Thesis for the Master's degree in Molecular Biosciences
Main field of study in Molecular Biology

Deletion of the nucleotides between + 54 and + 95 in the 5` UTR of the *Chlamydomonas* chloroplast gene *rbcL* and its effect on transcript stability

Ragni Ingeborg Monik Fet

60 study points

Department of Molecular Biosciences
Faculty of mathematics and natural sciences
UNIVERSITY OF OSLO 06/2007

ABSTRACT

In the 5` UTR of transcripts of the *Chlamydomonas* chloroplast gene *rbcL* secondary structures in the shape of two stem loops, one large and one small, have been identified. The large stem loop has been found to mediate folding of a ten nucleotides long sequence, stretching from position +38 to +47 relative to the transcription start site (TSS) (+1). This region has been found crucial for transcript stability. If its sequence, and/or partially its conformation, is disturbed, transcripts will be rendered unstable. Little focus has been aimed at the smaller loops function. In this study a deletion of nucleotides between positions +54 and +95 in the 5´ UTR of *rbcL* still renders stable transcripts, suggesting that the deleted region is not essential for transcript stability. This helps to narrow down the sequence and the number of nucleotides where a cis-acting element could be located, and to identify trans-acting factors believed to bind to this element, aiding in stabilizing the transcripts of the *C. reinhardtii* chloroplast gene *rbcL*.

PREFACE

This thesis is the final completion of a Master of Science degree at University of Oslo. The practical work was carried out in the laboratories of Professor Uwe Klein in the period from April 2006 to January 2007, while the work with writing up the thesis was done in the time between November 2006 and May 2007. This Masters degree in Molecular Biology has been completed at the Department of Molecular Biosciences where I have been a student from August 2005 to June 2007.

I would like to thank Professor Uwe Klein for letting me be a part of his group and for his guidance throughout this experiment. I would also like to thank Professor Jon Nissen-Meyer for taking the time to answer my questions and for giving me detailed information on possible future experimental approaches.
Last but not least I would like to thank my husband Chris for giving me encouragement throughout the entire process.

Ragni Ingeborg Monik Fet
Oslo, June 2007

ABBREVIATIONS

A:	Ampicillin
ATP:	Adenosine triphosphate
bp:	Base pair
C. reinhardtii:	*Chlamydomonas reinhardtii*
CFU:	Colony Forming Unit
CsCl:	Cesium Chloride
ER:	Endoplasmatic reticulum
e.g.	*Exempli gratia* / "for example"
et al.:	*Et alii* / "and others"
EtBr:	Ehidium Bromide
GTP:	Guanosine Triphosphate
GUS / *uidA*:	β-glucuronidase
HMW:	High Molecular Weight
HS:	High Salt
HSHA:	High Salt High Acetate
I:	Integration
IR:	Inverted repeats
kDa:	Kilo Dalton
L:	1 kb plus ladder
L:	Large subunit
LB:	Luria Bertani
mt:	Mating type
NaOH:	Sodium hydroxide
NEP:	Nuclear-Encoded Plastid Polymerase
NI:	No integration
PEP:	Plastid-Encoded Plastid Polymerase
PGA:	Phosphoglycerate
PNP:	Polynucleotide phosphorylase

pI:	Isoelectric point
pmol:	picomol
PSI:	Photo system I
PSII:	Photo system II
psi:	Pound-force per square inch
Rubisco:	Ribulose-1,5-biphosphate carboxylase / oxygenase
RuBP:	Ribulose-1,5-biphosphate
S:	Small subunit
SSC:	Sodium chloride-sodium citrate
TRP:	Tetratricopeptide Repeat
TSS:	Transcription Start Site
UTR:	Un-translated region
- “number” :	Indicates position in gene sequence upstream of TSS (+1)
+ “number”:	Indicates position in gene sequence downstream of TSS (+1)

CONTENTS

1 INTRODUCTION

1.1 Background

This study is part of a project aiming to narrow down the 5' UTR of the *Chlamydomonas* chloroplast gene *rbcL* as much as possible without affecting transcript stability. This is to find the minimum sequence and features required for transcript stability and for easier identification of cis-acting elements and trans-acting factors that presumably bind to these elements. To this date several modifications have been made to the 5`UTR of the *Chlamydomonas* chloroplast gene *rbcL*. To give the reader an idea of why a *deletion* was made, and why in that exact region, an overview of some of the previous constructs made is included (see Figure 1). Former students in Uwe Kleins group have also, as part of *their* Masters degree, modified the 5' UTR of *rbcL* in various ways, *e.g.* by adding both 10 and 20 nucleotides to its 5´ termini. The ten nucleotides addition (Zarins 2006) took on a single stranded conformation while the twenty nucleotides addition (Witsø 2006) folded into a loop structure. These additions both abolished accumulation of chimeric GUS transcripts. The structure of their projects as far as strains, media and methodology used were quite similar to this project.

Based on previous constructs made (see Figure 1), at least four things are clear:

- Any changes that will affect the sequence of the single stranded region between the two stem loops, from position +38 to +47 relative to the TSS at +1, will render unstable transcripts (see Figure 3 for details on that region).
- Any changes that affect the conformation of the first six nucleotides of this region (+38 to +43) will render unstable transcripts.
- A stem loop must be present in the place of the endogenous large loop since it is required to mediate correct folding of this single stranded region.
- No more than six nucleotides should be added to the 5´ termini of the *rbcL* 5´ UTR.

Ergo, a deletion was made that did not interfere with or contradict any of these statements.

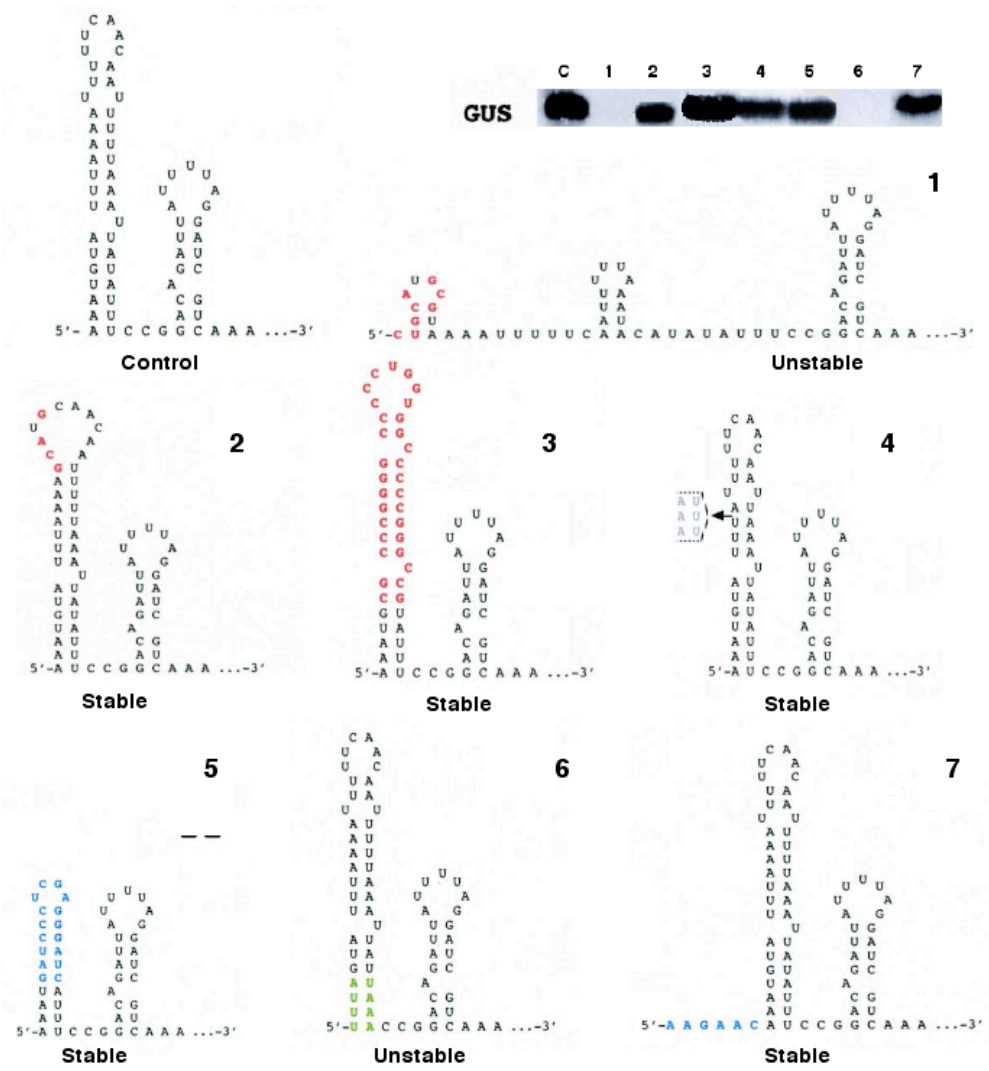

Figure 1: An overview over some of the chimeric gene constructs made by Suay et al. (2005), all showing the secondary structures of the 5` UTR. The construct in the top left corner, labeled as the control, is the endogenous 5` UTR of rbcL. Numbers 1 through 6 annotates the constructs. Red letters/bases indicate change of nucleotides; green indicates inversion of sequence; blue indicates an insertion of nucleotides, while dim grey indicates a deletion or replacement (only construct 11 where the sequence between +5 to +37 is replaced by a foreign sequence) of a sequence. The penal in the upper right corner shows the accumulation of GUS from each construct (C representing the endogenous rbcL 5´ UTR as a control). In the penal, a band under the number of the respective construct shows that transcripts accumulated, which means that the construct rendered stable transcripts. The varying intensities of the bands show the degree of stability compared to the band from the control, which represents complete stability. Modified after Suay et al. (2005).

1.2 Goal

Based on the background information provided in section 1.1, the nucleotides between position +54 and +95 (not including +54 but including +95), relative to the transcription start site (TSS) at position +1, in the 5`UTR of the *Chlamydomonas* chloroplast gene *rbcL* are deleted.

The goal is to find out if this deletion affects transcript stability.

The results of this study will help to further narrow down the number of nucleotides in the 5`UTR that are required for stability.

This modification deletes the sequence of almost the entire second stem loop, and contains a ribosome binding site important for translation. Therefore this deletion should also tell whether or not there is a link between translation and transcript stability in the *Chlamydomonas rbcL* gene.

1.3 Organization of the thesis

This thesis starts with an introduction of various promoters and the polymerases interacting with these. Further, a section describing RNA processing and the degradation of transcripts is included, since this is an important part of understanding the longevity of mRNA transcripts and which parameters affect this. It also aims to shed *some* light on the way expression of various genes depend on intercommunication between the nuclear and plastid genomes, and on how the mechanisms of gene regulation and the genetic makeup of *C. reinhardtii* compare to that of other organisms with similar genes. Some descriptions of previous research are included since these in part are what this entire experiment is founded upon.
To set these things into system, an introduction of the green algae *Chlamydomonas reinhardtii* is included, focusing on its versatility as a model organism, its physical abilities and its three genomes. These topics constitute chapter two of the thesis.

In chapter three a description of the materials and methods that were used to carry out this experiment follows. These include information on the various strains used and the

conditions under which these were grown, which vectors and plasmids were used, enzymes, media and more.

Chapter four focuses in detail on the actual results achieved, while in chapter five of the thesis the results are elaborately discussed. In the end of the discussion a series of future research approaches are suggested. These suggested approaches are based on information obtained through dialogues with professors that are proficient in the specific areas discussed and is not really within my field of research in this experiment. I still find it interesting and I want to include this to be able to give a more complete picture of the purpose of this experiment.

Chapter six presents conclusions based on the goals of this study and on the obtained results.

Chapters seven and eight include references to literature, and to the internet, respectively.

2 THEORETIC BACKGROUND

2.1 Transcription in the chloroplast

It is believed that chloroplasts originated from the primary endosymbiotic acquisition of oxygenic photosynthetic bacteria by non-photosynthetic cells. In *C. reinhardtii* many of the original bacterial genes are now present in the nuclear genome, where they have been integrated and stably maintained (see more about the chloroplast and its genome in sections 2.3.3 and 2.3.3.1). It may come as no surprise that of the three types of promoters that have been identified in chloroplast genomes, two resemble typical prokaryotic-like promoters rather than eukaryotic. One is an *E. coli* σ^{70} or σ^{43}-like promoter, comprising -10 (TATAAT) and -35 (TTGACA) consensus sequence elements. The other promoter, which has an unconventional structure, lacks the -35 element but has an extended -10 (TATAATAT) element. These have both been identified in the *C. reinhardtii* chloroplast genome. The latter is the most common protein gene promoter in the *Chlamydomonas* chloroplast while the former appears to be the more active one. A third promoter type identified in *Chlamydomonas* is believed to be an internal tRNA promoter (Salvador *et al.* 2004; Hajdukiewicz *et al.* 1997; Klein *et al.* 1992; Klein *et al.* 1994)

The basic *rbcL* promoter is located between -18 and +63, relative to the transcription start site (TSS) at +1. When this region was fused to the β-glucuronidase (GUS) encoding region of the *uidA* gene, the gene was transcribed only about 1-15 % as actively as the endogenous *rbcL* gene in chloroplast transformants. When the *rbcL* region included in the chimeric genes was extended to at least include the region up until +170, transcription rates were as high as for the endogenous *rbcL*. Thus, it was concluded that DNA sequences downstream of the basic *rbcL* promoter are required for endogenous level rates of transcription (Klein *et al.* 1994).

When putative promoter region sequences of eight *Chlamydomonas* chloroplast genes were compared, six of the sequences lacked a typical conserved -35 element, while seven of the sequences contained a palindromic TATAATAT motif in their -10 region (Figure 2) (Klein *et al.* 1992).

Figure 2: Comparison of putative promoter region sequences upstream of TSS in eight Chlamydomonas chloroplast genes. Six of the sequences lack~~ed~~ a typical conserved -35 element, while seven of the sequences contained a palindromic TATAATAT motif in their -10 region (boxed regions). This motif consists of two overlapping -10-like elements. Second to the bottom is the promoter of the rbcL gene. The promoter has nucleotides between positions -33 and -38 that are homologous to a -35 element in addition to the conserved -10 region. Note that the putative -35 region in the rbcL 5' region does not function as a promoter element. The numbers on the very top of the picture indicate positions relative to the TSS (+1), while the numbers below the sequence denote the position of this sequence in the original sequence report (Klein et al. 1992).

Genes in plastid genomes of higher plants are generally transcribed by two distinct RNA-polymerases; nuclear-encoded plastid polymerases (NEPs) and plastid-encoded plastid polymerases (PEPs). Both types of polymerases have been identified in photosynthetic higher plants and algae. Plastid genes can be divided into three classes: Those that contain only PEP promoters (class 1), those that contain only NEP promoters, (class 2) promoters and those that contain both PEP and NEP promoters (class three). Studies have shown that

genes coding for proteins of the two photosystems (PSI and PSII), *e.g., psbA,* and *psbD*, belong to class 1, while genes involved in genetic systems, *e.g atpB, atpI and clpP,* belong to class 3. Therefore it has been proposed that genes involved in the initiation and maintenance of the genetic systems during plastid development, *e.g. accD*, belong to class 2. In mature chloroplasts, transcription of these genes is likely largely replaced by PEP. Also, it appears that similarly functioning genes are transcribed by PEP, or by both PEP and NEP as a group. Transcription of all photosystem genes by PEP is a good example of this (Magee and Kavanagh 2002; Hajdukiewicz *et al.* 1997; Stern *et al.* 1997; Allison *et al.* 1996; Klein *et al.* 1994; Klein *et al.* 1992). As mentioned, a NEP has not been found in *Chlamydomonas*.

2.1 Processing and degradation of chloroplast mRNA, and mRNA stability

Chloroplast RNA processing and degradation are orchestrated by factors encoded by the nucleus (Rymarquis *et al.* 2006), and several nuclear mutations that affect the stability of chloroplast transcripts have been described. As examples of such mutations, the 6.2z5 and the GE2.10, described for *Chlamydomonas*, which destabilizes the products of the *psbB* and *psbC* genes, can be mentioned. These transcript-destabilizing mutations seem to be quite gene specific, each affecting only one or very few, chloroplast transcripts. This is contrary to what happens in higher plants, where related mutations can affect larger numbers of transcripts. Examples of such mutations are the *hcf109* nuclear mutation in *Arabidopsis*, which reduces the stability of transcripts from the *psbB, psbD, psbC*, and *ndhC* operons (Meurer et al. 1996).

RNA processing describes the process in which newly synthesized RNA molecules are modified. It has been suggested that mRNA molecules in the chloroplast likely are subject to many forms of processing, including formation of mature 5`- and 3` ends, intron splicing, RNA editing, polyadenylation and more. The machinery responsible for chloroplast RNA degradation has to date not been fully defined. Some processes have been proven to occur while others are merely suggested that they *may* happen.
Some factors that have been defined, and that carry out the most common reactions in which RNA is processed, are ribonucleases. These are ubiquitous enzymes found in bacteria, eukaryotes and organelles. Ribonucleases are the enzymes that ultimately carry out RNA maturation and degradation and they exist as two kinds: exoribonucleases, which

progressively remove nucleotides from transcripts in the direction of either 5´→3´or 3`→5`, and endoribonucleases, which cleave the transcripts internally. The products of the latter often serve as substrates for exoribonucleases (Monde *et al.* 2000). It has been suggested that a 5´→3´ exoribonuclease exists in the *Chlamydomonas* chloroplast, and the fact that such an exoribonuclease plays a role in transcript degradation has been suggested to be unique to *Chlamydomonas*. This came from the discovery that a *petD* mutant, *mcd1*, that could not accumulate *petD* transcripts, accumulated *petD* transcripts once a poly-G sequence was inserted into the *petD* 5´ UTR. Since poly-G sequences are known to obstruct the movement of 5´→ 3´ exonucleases it seems natural to conclude that a 5´→ 3´ exonuclease is involved in *Chlamydomonas* chloroplast mRNA degradation (Drager *et al.* 1998).

Chloroplast transcripts do not have a 5´ cap, whether they are primary transcripts or processing products derived from polycistronic precursors. Still, the 5´ ends of primary transcripts and processed transcripts are distinguishable. This is because the primary transcripts contain a terminal di- or triphosphate group that can be capped in vitro with labeled GTP and guanylyltransferase, whereas 5´ ends resulting from ribonucleolytic cleavage have free hydroxyl groups, which cannot be capped. In the spinach *psbB* mRNA for example, two 5´ ends have been observed; One corresponding to the transcription initiation site and one corresponding to a processing site. In *Chlamydomonas*, only transcripts with a processed 5´ end can be translated. This is consistent with the fact that mRNAs resulting from processing accumulate in *C. reinhardtii* (Monde *et al.* 2000; Nickelsen *et al.* 1999; Vaistij *et al.* 2000; Drager *et al.* 1998).

3` termini of most nuclear-encoded mRNAs are formed by endonucleolytic cleavage and polyadenylation. In contrast, prokaryotic and organellar mRNA 3` termini are formed either directly by transcription termination or by the processing activities of endo- and exoribonucleases (Stern and Kindle 1993). It was found that transcription termination accounted for at most 50% of the formation of 3`ends in *atpB* mRNA (Rott *et al.* 1996).

A factor that has shown to be of great importance in conferring mRNA stability is the formation of secondary structures in the 5` and 3` UTRs. Chloroplast protein-coding genes, like most bacterial genes, are generally flanked at their 3` ends by inverted repeats (IRs) that can fold into stem-loops (Rymarquis *et al.* 2006). In eukaryotes, elements important for mRNA longevity have been delineated primarily in the 3` end of transcripts, whereas in

prokaryote type mRNA (like bacterial- and organellar mRNA) the essential determinants seem to be mainly located in the 5` UTR (Suay *et al*. 2005). Several studies have focused on RNA secondary structures, *e.g*. stem loops, and their involvement in stabilizing transcripts, but their exact role in prolonging mRNA longevity is not clear. The *C. reinhardtii* chloroplast gene *rbcL* contains two stem loops, one large and one small, in its 5` UTR (Figure 3). It was previously believed that the stem loops kept the transcripts stable by preventing ribonucleases from binding to, and/or by impeding their movement along, the RNA. But a later study (Suay *et al*. 2005) showed that this was not the case. Instead, the large loop has a formative function in that it mediates folding of a ten nucleotides long sequence around its base into a specific conformation consisting of a helical and a single stranded region (see the boxed region in Figure 3). The common conclusion is that transcript degradation is accelerated significantly when changes are made to this ten nucleotides long region between positions +38 and +47, relative to the TSS at position +1 (see Figure 1 for examples). This was delineated as the real structure required for longevity of *rbcL* in chloroplasts (Suay *et al*. 2005).

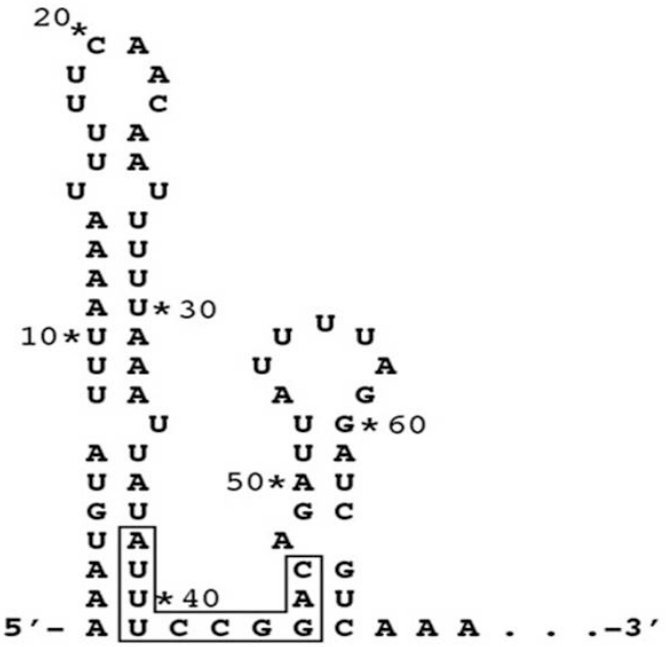

Figure 3: RNA secondary structure at the 5` UTR of the endogenous *C. reinhardtii chloroplast gene rbcL as predicted by RNA folding programs and verified in vivo by alkylation with dimethyl sulphate (Antonisen et al. 2001). The boxed region shows the 10 nt area between +38 and +47 that was delineated as the real structure required for transcript longevity in rbcL. From Suay et al. (2005)*

Such cis-acting sequences in the RNAs of genes encoded by the chloroplast have been found to be important determinants for mRNA longevity (Salvador *et al*. 1993). The same type of elements have been delineated both in the 5´ UTR and in coding regions of transcripts. As examples the *petD* (in the 5´ UTR) and the *atpA* genes (in the coding region) can be mentioned (Drapier *et al*. 2002; Sakamoto *et al*. 1992).

Trans-acting factors, either in the form of proteins or multi-subunit protein complexes, interact with these elements and confer stability upon the transcripts, likely by protecting the exposed 5` ends from nucleolytic attack.

In *C. reinhardtii*, three pleiotropic nuclear mutations, *mcd3, mcd4 and mcd5* (from here on referred to as *mcd3/4/5*) were identified. Originally these mutations were found to be photoautotrophic repressors of L2 and L6 p*etD* mutants. L2 and L6 were delineated as translationally defect, transcript-destabilizing mutants of p*etD*, and accumulated only 1-3 % of wt p*etD* mRNA levels. Also, no cytochrome b_6/f complex subunit IV, which is encoded for by the p*etD* gene and required for photosynthesis, was generated. Through analyses of thirty two genes they found that mcd3/4/5, likely due to their pleiotropic nature, prevent degradation of transcripts of seventeen of these thirty two genes, from five different gene clusters. This indicates that MCD3/4/5 are part of several multi-protein complexes that bind to regulatory regions, and are responsible for the maturation and degradation of transcripts in *C. reinhardtii* chloroplasts (Figure 4) (Rymarquis *et al*. 2006).

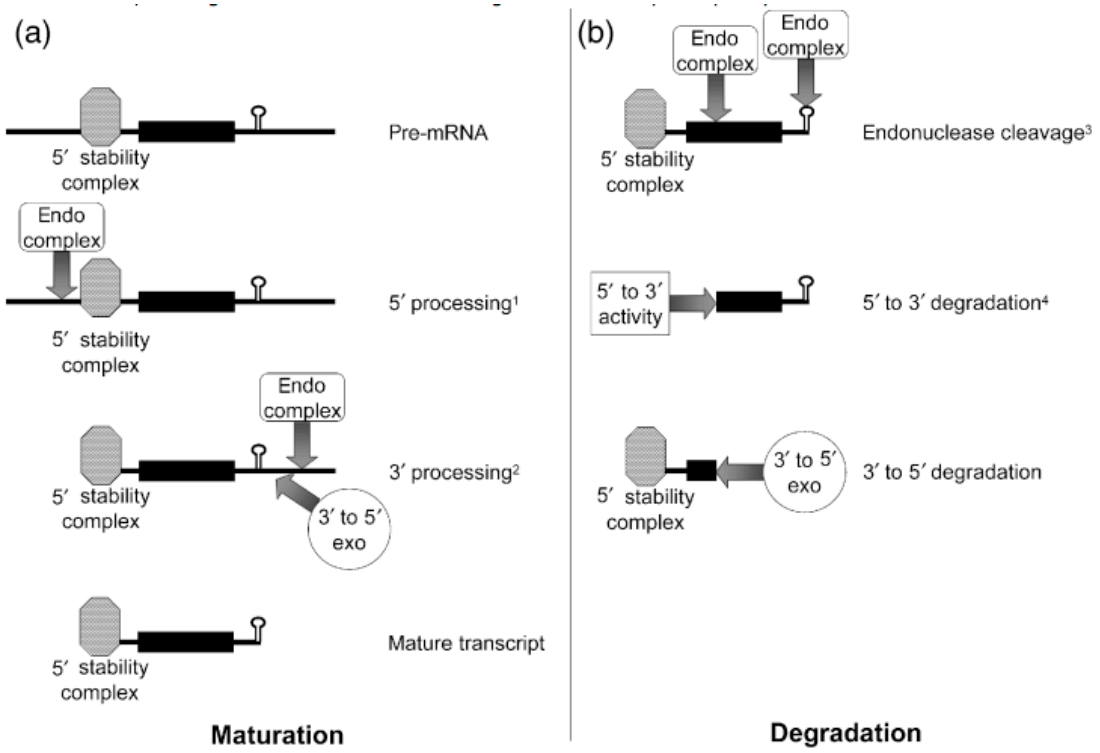

Figure 4: A working model of chloroplast maturation and degradation pathways in Chlamydomonas. (a) The steps involved in RNA maturation for a generic RNA are shown here, not necessarily in the order of occurrence. Coding regions are shown as black boxes, and inverted repeats known to stabilize the 3' ends of transcripts are shown as stem-loops. MCD/4/5 are predicted to be part of the endonucleolytic complex (endo complex) involved in both 5' and 3' processing. In 3' processing, the endonuclease cleavage is followed by 3' to 5' exonucleolytic trimming ("3' to 5' exo"), as it occurs for atpB (Stern and Kindle 1993). (b) The same endo complex as in ***a*** *could also participate in degradation initiation, leading to or stimulating the 5' to 3' pathway involving a 5' to 3' exonuclease or multiple endonuclease cleavages by the endo complex. The 3' to 5' pathway is similar to 3' processing, except that the secondary structure has been removed, allowing complete degradation of the transcript by 3' to 5' exonucleases. (Rymarquis et al. 2006.)*

There are considerable experimental *indications* that trans-acting factors, in the form of transcript-specific RNA binding proteins encoded by the nucleus, protect RNAs from nucleolytic attack. As an example, the nuclear *Nac2* gene, which encodes an mRNA-stabilizing protein, is required for transcripts of the chloroplast *psbD* gene to be stable. Through UV cross-linking studies, a 47 kDa protein was found to bind to the *psbD* 5´UTR.

When this binding activity was altered, a marked loss of stability was observed for *psbD* mRNA. *Nac2* cDNA was isolated and found to encode a novel tetratricopeptide repeat (TRP) – containing protein. The C-terminal end of *Nac2* was found to contain nine TRP-like domains, while the N-terminal end, on the other hand, appeared to be non-essential for its function. TRP proteins are known to be involved in various biological processes and have been found to often function as scaffolding proteins aiding in assembly of other proteins into multi-subunit complexes. At least one of these subunits was found to be essential for proper folding of *Nac2*. *Nac2* also appears to be part of a high molecular weight (HMW) complex associated with RNA. Another gene, *Mbb1*, has been found to restore *psbB* transcript stability in nuclear mutants that did not accumulate *psbB* transcripts. The *Mbb1* gene product was localized to the chloropast stroma but has not been shown to associate with *psbB* mRNAs (Vaistij *et al*. 2000). *Mbb1* has also been found to contain TRP domains and is also part of a HMW complex. Based on these findings it seems likely that each HMW complex specifically recognizes one chloroplast 5´ UTR and recruits a common factor through its TRP. Such a factor could be a nuclease, a translation factor or a processing enzyme. This could be a common mode of regulation of all chloroplast genes, including *rbcL*. Still, binding of the *Nac2* protein to RNA has not been demonstrated to date (Boudreau *et al*. 2000).

In *E. coli* on the other hand it appears that the mere presence of a stem-loop in the immediate vicinity of the 5´ terminus is crucial for stability, whereas the sequence of the stem-loop is quite unimportant. Thus, the lifetime of a normally labile mRNA can easily be prolonged in *E. coli*, just by adding a simple hairpin structure at its 5´ terminus (Emory *et al*. 1992).

Stem loops in the 3´ ends have also been suggested to play a role in stabilizing transcripts. As the case initially was for the 5´ UTRs, it has been proposed that the actual 3´ secondary structures themselves protect transcripts from degradation. The mechanism(s) for protection have not been defined but it is possible that the 3´ ends are protected by specific trans-acting factors or by similar features (Stern and Kindle1993).

Several other factors have also been both suggested and/or shown to have an effect on transcript stability and degradation. Amongst these is the redox state. Expression of chloroplast genes is under control of both external factors and internal processes. Light is the most important external factor implicated in chloroplast gene expression since it affects transcription, transcript stability, translation, protein modification and turnover of proteins.

Light initiates electron transport in chloroplast, and although it is a little unclear how, it has been suggested that electron transport does have an effect on the processes mentioned above. Research indicates that redox carriers transfer signals between electron transport and various molecular processes. Investigation of the *Chlamydomonas* chloroplast gene *psbA* revealed that redox signals likely were being carried by several thiol proteins in the electron transport chain to a multi-protein complex that activates translation by binding specifically to the *psbA* transcript (Salvador and Klein 1999). When isolating RNA from cells harboring chimeric *rbcL:*GUS gene construct they are often grown in alternating 12 hr light / 12 hr dark cycles and RNA is isolated towards the end of a dark period. It has previously been shown that chimeric *rbcL:*GUS transcript accumulation from RNA isolated at the end of a dark period compared to the end of a light period is much higher due to more stable transcripts. By comparisons, transcripts of the endogenous *Chlamydomonas* chloroplast *rbcL* gene are as good as equally abundant at all times when grown in alternating light and dark cycles. This is because the endogenous *Chlamydomonas rbcL* harbors, between nucleotides +21 and +41, a target for photo-accelerated degradation. Simultaneously, it has sequences between nucleotides +14 and +27, and +329 and +334, which counteract this effect by interacting with proteins that in turn can physically block the target for photo-accelerated degradation from nucleolytic attack (Singh *et al.* 2000). In summary: chimeric genes involving only the 5` UTR of the *rbcL* gene include the target for photo-accelerated degradation but not the two protecting elements. Therefore, transcripts from chimeric *rbcL:*GUS genes are less stable in the light than transcripts from the endogenous *rbcL* gene (Salvador *et al.* 1993; Singh *et al.* 2000). Although the factor or factors responsible for conferring light-mediated regulation on the *rbcL* gene are unknown, it is *controlled* by the redox state of the chloroplast (Salvador and Klein 1999). Thus, also a light-regulated pathway for degradation of chloroplast *rbcL* transcripts exists in *Chlamydomonas*.

As shortly mentioned earlier, polyadenylation has also been proposed as a factor important in degradation of RNA. Komine *et al.* (2000) suggested that polyadenylation of pre-mRNA competes with other RNA processing pathways. Prior to this there had been found no evidence that polyadenylation occurred in *Chlamydomonas* chloroplasts, and other similar studies had only found polyadenylation to occur on mRNAs in other chloroplasts and in bacteria. This study showed that polyadenylation *does* occur in the *Chlamydomonas* chloroplast and that it affects not only the stability of mRNAs but also the stability of tRNAs and rRNAs (Figure 5). They also reported that most of the poly(A) tails are added to internal sites of the RNAs, suggesting that these tails often are added to either incompletely synthesized and/or to partially degraded transcripts (Komine *et al.* 2000).

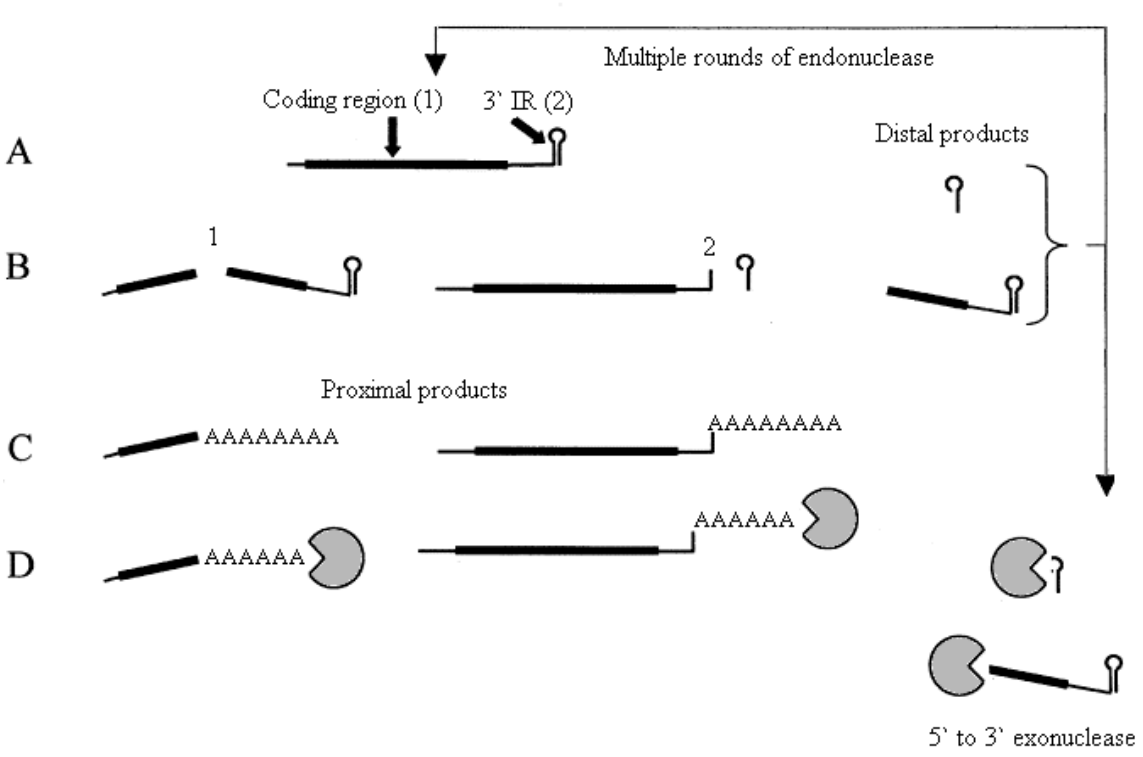

Figure 5: A suggested working model for chloroplast mRNA degradation where polyadenylation is involved. The scheme is shown for a typical chloroplast mRNA with a 3' IR, such as Chlamydomonas atpB. ***A)*** *Initially, endonucleolytic attack occurs within the coding region (1) or the 3' IR (2), catalyzed by various enzymes.* ***B)*** *This cleavage yields proximal and distal products. The distal products are subject to further rounds of endonucleolytic cleavage, or may be degraded by a 5' to 3' exonuclease activity.* ***C)*** *The proximal products are efficiently polyadenylated with a tail up to several hundred nucleotides in length.* ***D)*** *The polyadenylated RNA molecule is rapidly degraded by exonuclease(s), such as PNP (polynucleotide phosphorylase). Modified after Monde et al. (2000).*

Although this model is interesting, it seems strange that endonucleolytic attack would occur internally in the coding region. Some mRNAs are long lived, often with longevity of several hours, and such an internal cleavage would result in immediate degradation of the mRNA. Maybe such polyadenylation is aimed at specific mRNAs that are targeted for immediate degradation.

As mentioned, additional forms of RNA processing include intron-splicing and RNA editing.

Introns can be cis-spliced or trans-spliced. To date, three *Chlamydomonas* chloroplast genes have been found to contain introns. *In vivo* splicing has never been observed, but it is believed that the process requires trans-acting factors, including ribonucleases. (Rochaix 1996; Monde *et al*. 2000).

RNA editing in chloroplasts involve C to U base changes which normally leads to the formation of start codons or to the re-introduction of conserved amino acids. Many chloroplast mRNAs undergo RNA editing but there has been found no evidence for this occurring in *Chlamydomonas* chloroplasts (Monde *et al*. 2000; Maier *et al*. 1996).

2.2 Chlamydomonas reinhardtii

2.2.1 Characteristics of Chlamydomonas

Chlamydomonas is a large genus of the *Chlorophyte* division of algal protists, usually unicellular, with over 600 identified species. These algae are highly adaptive and have been found to live in many different environments throughout the world. They have been isolated not only from fresh water and soils, but also from marine waters and even snow (Adams 2000).

All species of Chlamydomonas share a basic body plan, while cell shape varies greatly among species. Most species are ellipsoid or ovate in shape, and have a clearly polar structure with two anterior flagella and a single basal chloroplast. All species have a distinct seven-layer cell wall, primarily consisting of hydroxyproline-rich glycoproteins, usually closely appressed to the plasma membrane. The nucleus is normally centrally located, with a prominent nucleolus. The nuclear membrane is continuous with the endoplasmatic reticulum (ER) (Hajdukiewicz et al 1997), and one to four Golgi bodies are situated nearby. All but a few species have a colored “eyespot” that senses light, and most have one or two contractile vacuoles that help control the osmotic pressure within the cells (Figure 6) (Harris 2001).

Chlamydomonas cells harbor three genomes; one in the nucleus, and two, "gene-poor", polyploid organelle genomes, one in the chloroplast and one in the mitochondria (Rochaix 2002; Kathir et al. 2003). The nuclear genome is about 100 Mb in size, and is divided into 17 chromosomes of unknown sizes (Kathir et al. 2003). The chloroplast genome is about 200 kb of size, while the mitochondrial genome is only 15.8 kb (Vahrenholz et al. 1993). Mitochondria are dispersed throughout the cytosol, and, due to their small genome, contain few genes (Harris 2001). The discovery of a large number of nuclear loci involved in chloroplast gene expression has led to the belief that genetic interactions between all three genomes exist (Rochaix 2002).

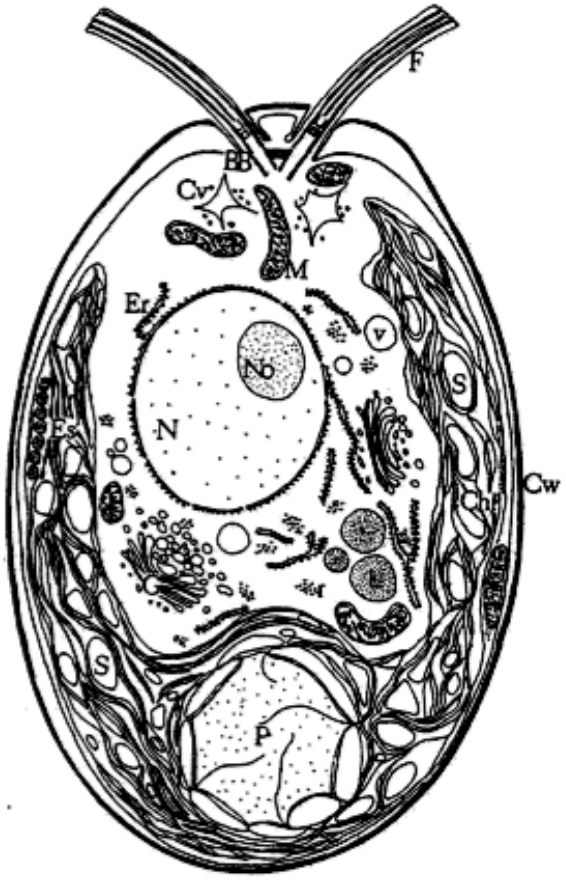

Figure 6: A semi-diagrammatic representation of an interphase Chlamydomonas cell. Cell length; 10 μm; BB, basal bodies; Chl; chloroplast; Cv, contractile vacuole; Cw, cell wall; Er, Endoplasmatic reticulum; Es, eyespot; F, flagella; G, Golgi apparatus; L, lipid body; Mi, mitochondria; N, nucleus; P, pyrenoid; r, ribosomes; S, starch grain; v, vacuole. From the Chlamydomonas Sourcebook (Harris 1988), originally by courtesy of John Harper.

2.2.2 The sexual cycle of Chlamydomonas reinhardtii

Chlamydomonas species average about 10 μm in diameter, with significant variations throughout the cell cycle. In mitotic growth the cells double as fast as every eight hours and are cheap and easy to maintain. Under periods of stress, such as nitrogen starvation, *Chlamydomonas* cells differentiate into isogametes.

Two distinct mating types (mt) designated mt+ and mt- exist. These fuse sexually, thereby generating a thick walled diploid zygote. The zygote is not flagellated and it serves as a dormant form of the species in soil. After several days the diploid zygotic nucleus divides meiotically and releases four flagellated haploid cells that can resume the vegetative life cycle (Figure 7). The products can be separated, allowing tetrad analysis, or be mass scored by re-spreading onto fresh plates. A curious fact is that under ideal growth conditions, cells may undergo two or three rounds of mitosis before the daughter cells are released from the old cell wall and into the medium. Thus, a single growth step may result in four or eight daughter cells per mother cell. (Adams 2000; Harris 2001).

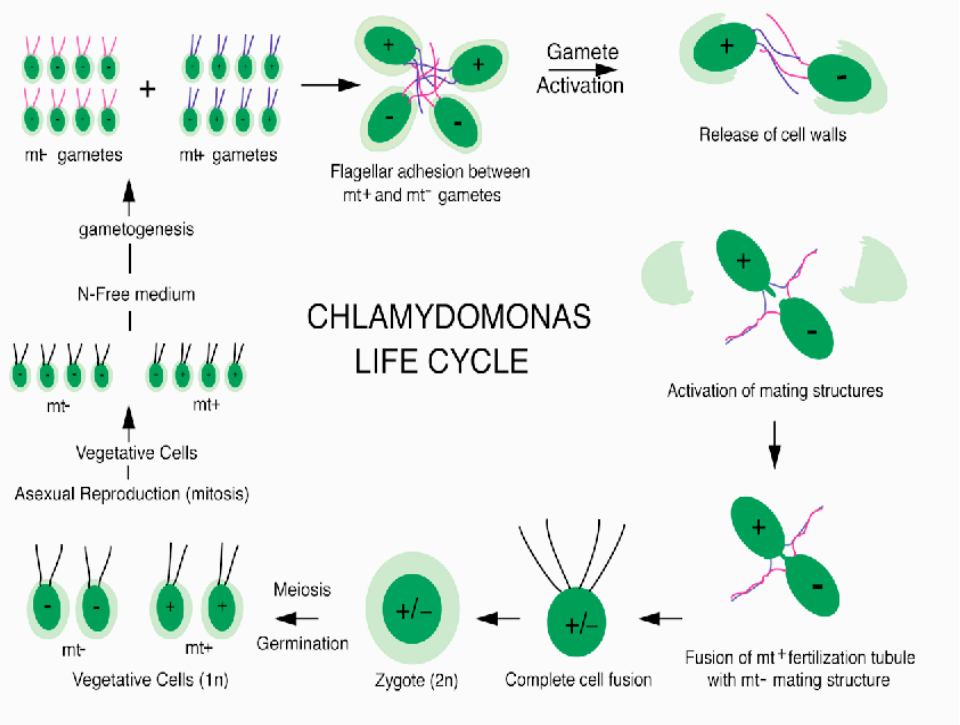

Figure 7: The sexual cycle of Chlamydomonas reinhardtii. From Harris (2001).

2.2.3 The chloroplast of Chlamydomonas reinhardtii.

2.2.3.1 The chloroplast genome

C. reinhardtii harbor a large cup-shaped chloroplast in their basal end that partially surrounds the nucleus, and, as mentioned earlier, it has its own genome. This was discovered over 40 years ago and has led to intense studies of both its structure and organization (Palmer 1985). Complete sequencing of the chloroplast genome of *C. reinhardtii* has revealed a map of 203 395 bp that possesses two copies of an inverted repeat sequence which harbors rRNA genes. It can exist in circular or linear forms, either as multiple circles with sizes of ca. 200 kb (Rochaix 2002) or as multiple small, linear DNA fibers of <100 kb in size (Maul *et al.* 2002). The genome is found to contain over 100 genes (Rochaix 2002) and nearly all structural genes encoding chloroplast components that are found in land plants have also been identified in *Chlamydomonas* (Harris 2001) (Figure 8). Sequencing of plastid genomes has also revealed potential RNA editing sites, although, as mentioned, not in *Chlamydomonas* (Monde *et al.* 2000), and co-regulation of genes or gene clusters (Maul *et al.*2002).

As mentioned in section 2.1., it is believed that chloroplasts have an endosymbiotic origin. Although there is still some controversy over whether all chloroplasts are ultimately derived from one single, or from several, endosymbiotic processes, it is clear that the evolution of the organelle involved the transfer of a large portion of the genetic information of the original endosymbiont(-s) to the nucleus of the host organism. Despite the fact that chloroplasts are found to contain more than a thousand different protein species, their genomes only encode about 100–200 proteins. The remainder of the chloroplast proteins are encoded in the nucleus and imported post-translationally through machinery that is partly derived from the chloroplasts bacterial ancestor. In higher plants two-thirds of the 60 or so chloroplast ribosomal proteins are encoded by the nucleus, even though these genes have a clear bacterial ancestry (Alberts *et al.* 2002).

Several suggestions have been made as to how transfer of genes to the nucleus might provide a selective advantage. Some have suggested that transfer to the nucleus protects the genes from oxygen free radicals, which may be generated during photosynthesis, while others believe that placing a gene in a sexual population (as opposed to the asexual population represented by the uniparentally inherited chloroplast) is advantageous. A third theory has been promoted; suggesting that movement of a gene to the nucleus avoids

deleterious effects due to the biased nucleotide composition associated with many chloroplast genomes. Another question that has been raised focuses on why a subset of genes has apparently resisted the selective advantages of chloroplast-to-nucleus transfer. Currently, not much is known of the molecular mechanism of transfer of genes from chloroplast (Lister *et al*. 2003; Maul *et al*. 2002; Palmer 2003).

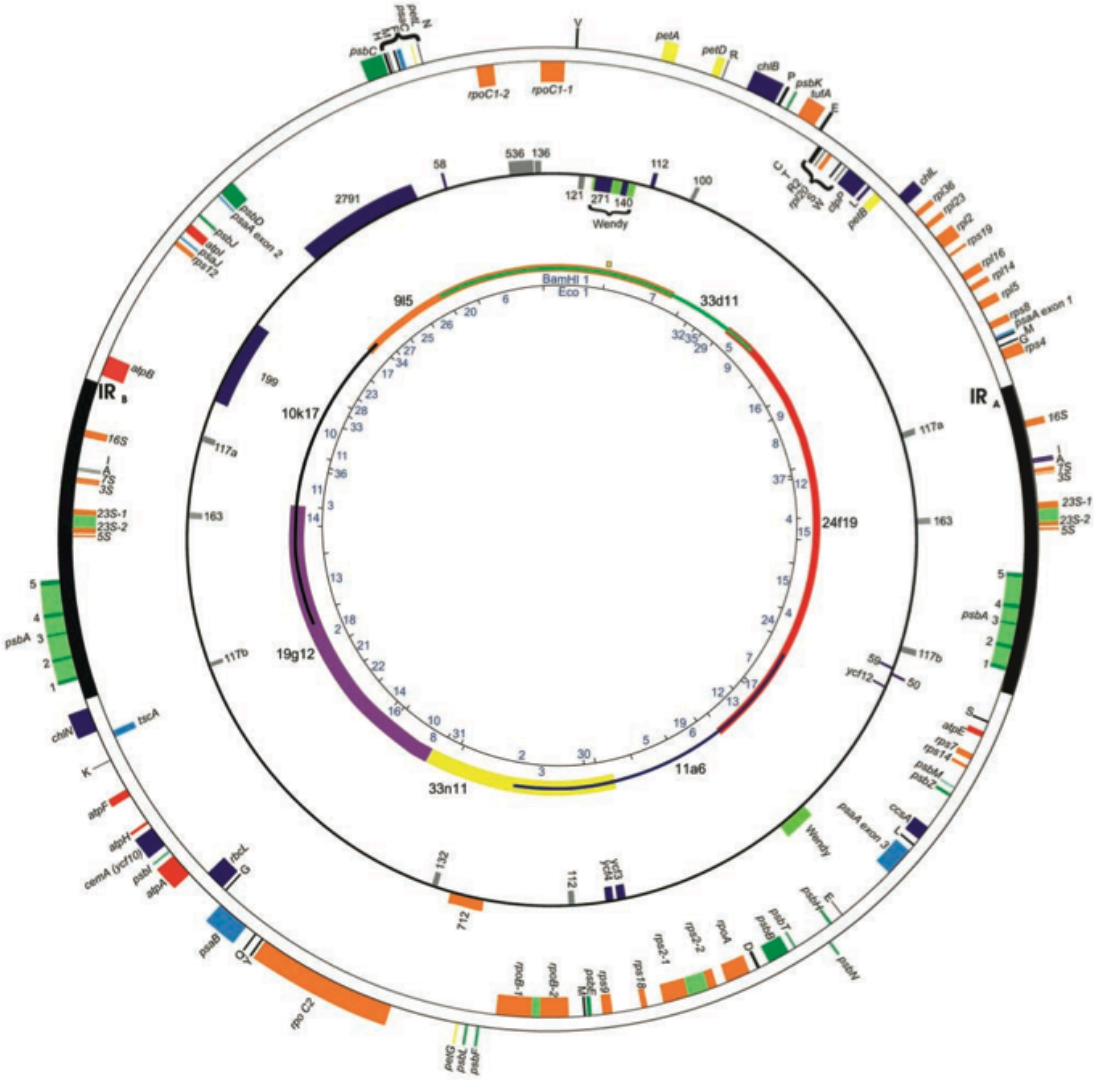

Figure 8: The plastid chromosome of C. reinhardtii. The inner circle shows BamHI and EcoRI restriction fragments mapped according to Rochaix (1980). The middle circle shows the seven overlapping BAC clones that span the genome, while the outer circle shows genes of known and unknown function. The black boxes on the right and left side (opposite each other) of the outer circle represent the two inverted repeats. From Maul et.al. (2002)

2.2.3.2 The chloroplast structure

The internal structure of the *Chlamydomonas reinhardtii* chloroplast is mainly made up of thylakoid membranes, bi-layered compartments consisting of phospholipids. They are interconnected and tend to form stacks of vesicles or discs referred to as granum (Figure 9). The chloroplasts envelope consist of a highly permeable outer membrane, a much less permeable inner membrane, in which membrane transport proteins are embedded, and a narrow inter-membrane space in between. The inner membrane surrounds a space called the stroma, which contains numerous metabolic enzymes. Since the chloroplast contains its own genome, the stroma contains a special set of ribosomes, RNAs and of course the chloroplast DNA (Alberts *et al*. 2002; Harris 1988). The broad basal area of the chloroplast contains a prominent pyrenoid. Chloroplasts are also the site of photosynthesis.

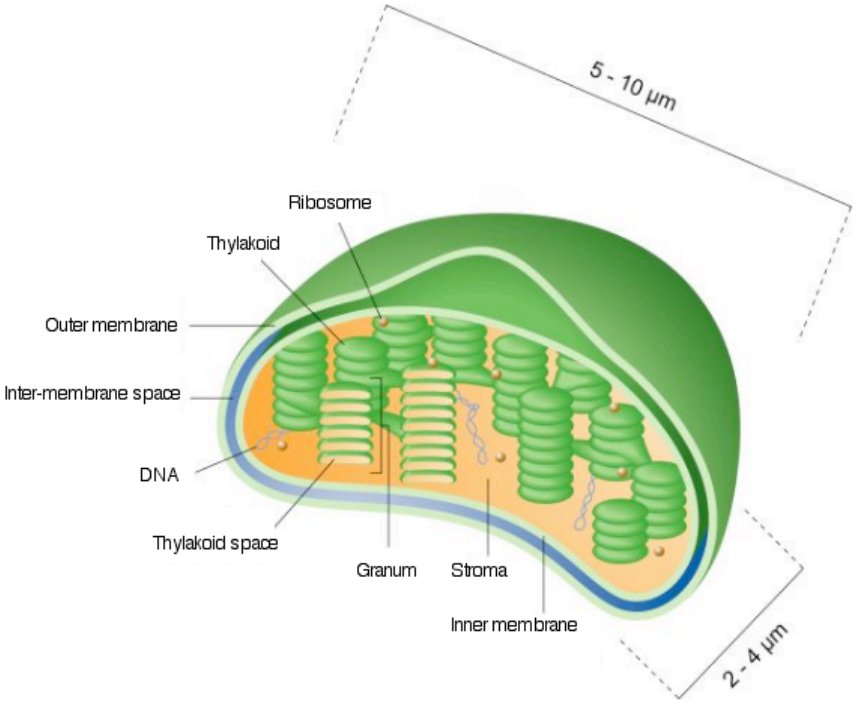

Figure 9: Illustrated cross section of chloroplast. A Chloroplast is about 5 – 10 µm long and has a total width of about 4 – 8 µm. Modified from www.vscht.cz (see section 8 for full web address).

2.2.4 Photosynthesis in Chlamydomonas reinhardtii

A phototrophic way of life is possible only for organisms that possess photoreceptors, photopigments, the photosynthetic electron transport chain, and a biochemical pathway for carbon reduction (Hudock and Levine 1964).

Unlike the mitochondria, the inner membrane of the chloroplast is not folded into cristae and does not contain electron transport chains. Instead, as shortly mentioned in the previous section, all the energy-yielding systems of the chloroplast, including its chlorophyll, are located in the thylakoid membrane.

Amongst the numerous metabolic enzymes found in the chloroplast stroma, or more specifically in the pyrenoid, is Ribulose-1,5-bisphosphate carboxylase/oxygenase, also known as Rubisco, the key enzyme in photosynthetic carbon fixation (Alberts *et al*. 2002; Harris 1988).

All the photosynthetic reactions that occur inside phototrophic organisms can be divided into two main categories; the "light reactions", or photosynthetic electron-transfer reactions, and the "dark reactions", also known as the carbon-fixation reactions (Alberts *et al*. 2002). In the "light reactions" an electron in the chlorophyll molecule of photosystem I (PS I) is, when sunlight hits it, excited to a much more electronegative state. This enables the electron to move through the electron-transport chain in the thylakoid membrane. Algae, *e.g*. *C. reinhardtii*, cyanobacteria and green plants obtain these electrons from H_2O, producing O_2 as a by-product. During the electron transport, H^+ is pumped across the thylakoid membrane, creating a proton motive force, which drives the synthesis of ATP, via an ATPase, in the stroma. In the final step of these reactions, high-energy electrons, together with H^+, are loaded onto $NADP^+$ converting it to NADPH. All these reactions happen in the chloroplast (Hudock and Levine1964; Alberts *et al*. 2002).

To drive the conversion of CO_2 to carbohydrates, the "dark reactions", the NADPH and the ATP produced in the "light reactions" serve as sources of reducing power and energy, respectively. The "dark reactions" begin in the chloroplast stroma and continue into the cytosol where they produce sugars and other organic molecules.

The "light– and dark reactions" are separate processes, the first requiring light at all times while the latter only requires it indirectly, but are interconnected by elaborate feedback mechanisms (Alberts *et al*. 2002).

2.2.4.1 Ribulose-1,5-bisphosphate carboxylase/oxygenase – Rubisco

Rubisco is a large enzyme, with a molecular mass of about 560 kDa. Nature elaborates two architecturally distinct, but functionally analogous, forms of Rubisco. The L_2-form; a homodimer of 50 kDa subunits found in purple, non-sulfur, bacteria, and the L_8S_8-form; a hexadecamer consisting of eight large 53 kDa subunits (L), and eight small 14 kDa subunits (S), hence the name. The L_8S_8-form is present in all other photosynthetic organisms, e.g. *C. reinhardtii* (Hartman and Harpel 1994; Spreitzer 1993).
The large subunits of L_8S_8 Rubisco, each of which contains an α/β-barrel active site for photosynthetic CO_2 fixation (Spreitzer 1993), are encoded by the chloroplast *rbcL* gene and synthesized on the chloroplast ribosomes. The small subunits are encoded by the nuclear *rbcS* gene family and synthesized in the cytosol (Yosef *et al.* 2004). After translation, newly synthesized small subunits are translocated across the chloroplast membrane, where an N-terminal signal peptide is proteolyzed prior to assembly with the large subunits (Spreitzer 1993; Hartman and Harpel 1994). As mentioned above, the large subunits harbor the actual active site for CO_2 fixation so their roles in this process are quite defined, while the small subunits` functions in these reactions have been more elusive. Spreitzer (2003) suggested that they have two roles; they help assemble and concentrate the large subunits, and they are in part responsible for the higher CO_2/O_2 specificity that L_8S_8-form Rubisco have compared to L_2-form Rubisco. Since the two types of subunits are coded for by two separate genomes, the synthesis and assembly of Rubisco require intercommunication between the nucleus and the chloroplast (Hartman and Harpel 1994; Spreitzer 2003).

Each molecule of the Rubisco complex works rather slowly, processing only about 3 molecules of substrate each second, so many enzyme molecules are needed. Because of this, Rubisco can constitute about 50 % of the total chloroplast protein, and it is by many believed to be the most abundant protein on earth.
Within the chloroplast Rubisco catalyzes the initial steps of photosynthetic carbon reduction, a cycle also referred to as the Calvin cycle, by combining CO_2 with ribulose-1,5-bisphosphate (RuBP) (Mizohata *et al.* 2002) to form two molecules of phosphoglycerate (PGA). However, O_2 competes with CO_2 at the same active site of Rubisco. Thus, oxygenation of RuBP leads to the production of only one PGA molecule, with the remaining two carbons from RuBP forming phosphoglycolate, which is the initial reactant of photorespiration (Figure 10). Photorespiration is a nonessential pathway that causes loss of CO_2. Since oxygenation and photorespiration are responsible for losing more than 30%

of the CO_2 that could potentially be fixed during photosynthesis, the question whether plant productivity could be improved by reducing oxygenase activity or increasing carboxylase activity has been raised (Spreitzer 1993; Spreitzer 2003)

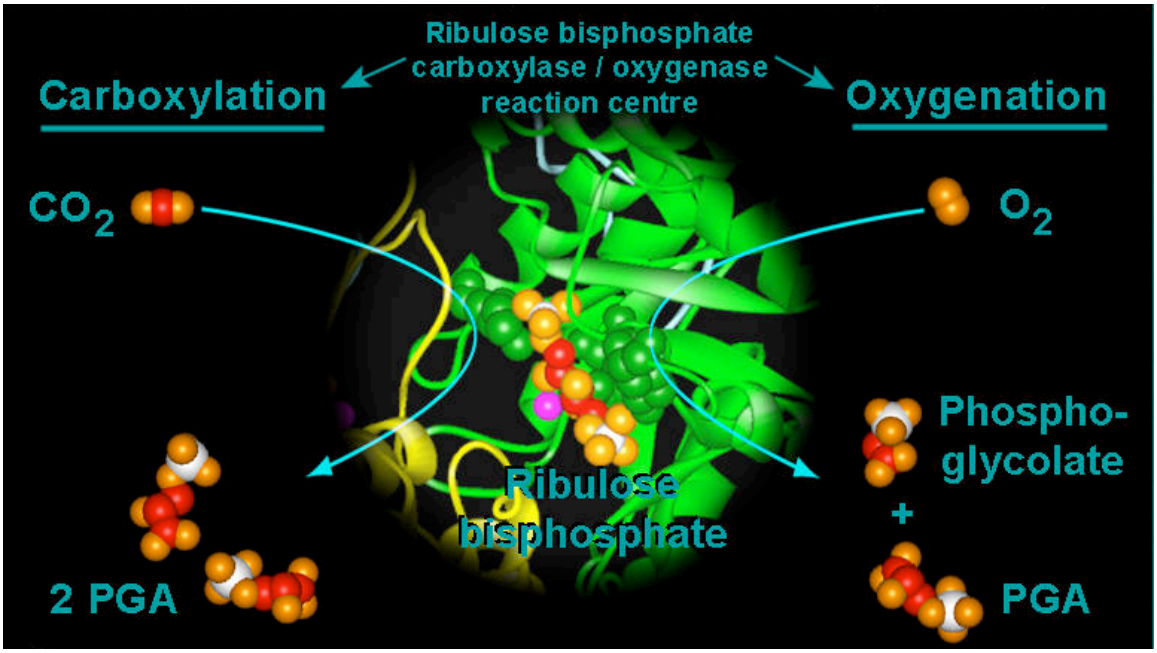

Figure 10: The active site of Rubisco. The active site of Rubisco cannot fully discriminate between carbon dioxide and oxygen, and so it catalyses reactions of carboxylation (producing 2 PGA molecules) and wasteful oxygenation (producing only one PGA to enter the Calvin cycle, plus one phosphoglycolate). From http://www.rsbs.anu.edu.au (see section 8 for full web address)

2.2.5 Chlamydomonas reinhardtii as a model organism

While the genus *Chlamydomonas* contains several species that have become popular as research tools, *C. reinhardtii* is by far the most frequently used. The algae has several features that makes it a preferable model organism, like its short generation time, easy maintenance, low cost and various mutant strains available (Harris 2001). But primarily it owes its success to its ability to grow non-photosynthetically with acetate as its sole carbon

source. It is believed that the principal laboratory strains of *C. reinhardtii* derive from isolates made by GM Smith from soil collected near Amherst, Massachusetts in 1945.
Recently, regulation of gene expression in chloroplasts has received a lot of attention. Scientists have found *C. reinhardtii* to be especially useful for these studies since it is amenable to both biochemical, genetic and molecular analyses. It has reliable and stable nuclear- and chloroplast transformation systems, something that has further broadened its versatility as a model organism (Stampacchia *et al.*1997).

Vegetative cells of the *C. reinhardtii* species are haploid with 17 small chromosomes and the nuclear genome has a high GC-content, approximately 62%. This can create difficulties when cloning genes, but *C. reinhardtii´s* other benefits as a model organism overshadows this drawback (Harris 2001).

3 MATERIALS AND METHODS

3.1 Strains and media

All cloning steps were carried out in the recombinant deficient *E. coli* strain TB1, an hsdR- derivative (restriction minus) of the *E. coli* strain JM83. Bacterial cells were grown on solid LB (Luria Bertani) agar plates or in liquid LB, containing 60 μg/ml ampicillin (LB+A), over night at 37 °C. The agar plates were initially made with ampicillin and stored at 4 °C for up to several weeks, but, since ampicillin is not stable in solution, liquid LB + A medium had to be prepared right before use and not kept for more than 1-2 weeks refrigerated.

The non-photosynthetic mutant strain ac-uc-2-21 mt+ (*atpB* mutant CC-373, from here on referred to as CC-373), obtained from the *Chlamydomonas* Genetics Centre at Duke University (Durham, NC), was used for transformation of the *Chlamydomonas* chloroplast. CC-373 is a non-reverting, acetate requiring and ATP synthase deficient mutant, with a deletion in the chloroplast *atpB* gene. The mutation involves a deletion of a 2.5 kb region comprising the 3´ end of the chloroplast gene *atpB*, and part of the 5` end inverted repeats (IR) (Figure 11).

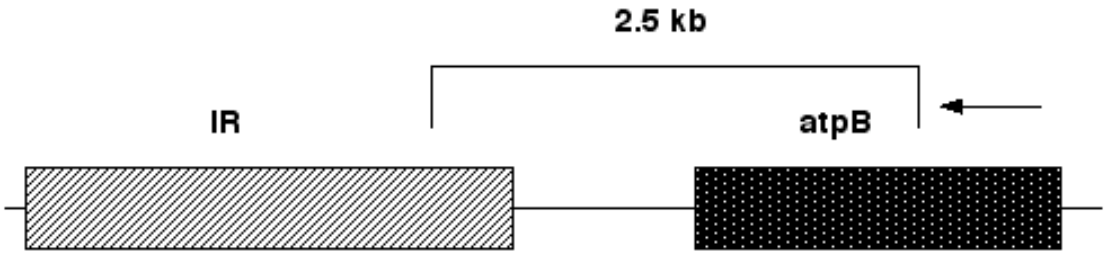

Figure 11: Deletion in the mutant CC-373. The lighter bar represents the inverted repeat (IR) region while the dark bar represents the atpB gene. In the mutant Chlamydomonas strain CC-373, a 2.5 kb large region is removed, deleting parts of both the atpB gene and the IR. This makes the mutant non-photosynthetic. The arrow indicates the direction of transcription. Drawing is not to scale.

Once the chimeric gene construct has been transformed into the chloroplast genome (see Figure 16 and Figure 17 for details) both the *atpB* gene and the IR will be completed by homologous recombination and photosynthesis is restored.

CC-373 cells were grown in High Salt High Acetate (HSHA) medium, made up from High Salt (HS) medium (Sueoka 1960) supplied with 2.5 g/l potassium acetate as 100 ml cultures, and maintained under low light with regular addition of fresh HSHA medium. This was done to keep the cultures at log phase density. The cultures were constantly supplied with air and CO_2. The cells were plated onto HSHA plates and kept at low light until the moment of transformation.

Transformants were kept in HS medium in Erlenmeyer flasks, under light at room temperature, as stock cultures.

The *Chlamydomonas* chloroplast transformants +157 and MU7 were used as controls for quantification of GUS abundance in the Northern blot. MU7 has an *rbcL:*GUS reporter gene in its chloroplast genome, located between its *atpB* gene and the IR (Figure 12). More detailed information on +157 and MU7 will follow in section 4.4.

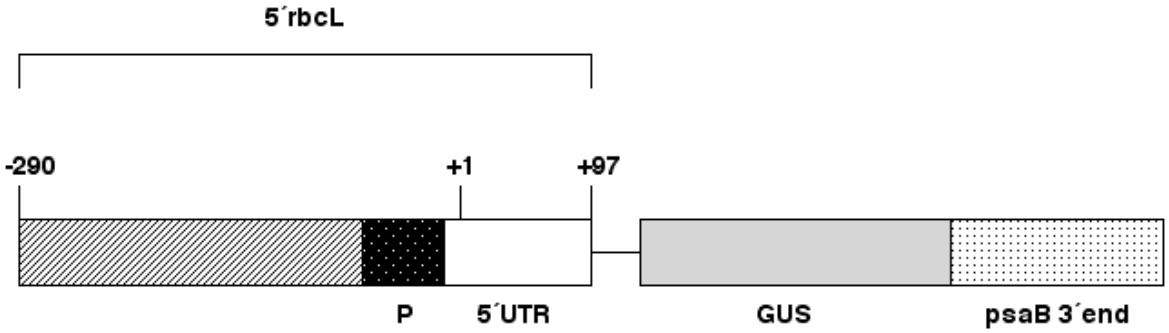

Figure 12: The MU7 construct. It consists of bases -290 to +97 of the rbcL region fused 3´end to the E. coli uidA gene (GUS). MU7 does not contain the enhancer region located around position +140 in the rbcL gene, but it harbors the RNA stabilizing element and the putative rbcL promoter (P). Drawing is not to scale.

3.2 Construction of pRF+54

3.2.1 Prediction of secondary structure

To be in accordance with the aim of investigation in this study, it was essential that the ten nucleotides long sequence between +38 and +47, located between the endogenous large- and small stem loop, remained undisturbed. At least as far as the sequence goes, and partly regarding the conformation (see sections 1.1 and 2.1 for details). It was also of importance that the endogenous large stem loop not necessarily remained undisturbed but that a secondary structure in the form of a stem loop formed in its place. As long as a loop mediated correct folding of the +38 to +47 region, its size and shape was insignificant. Although, the modification introduced to the 5` UTR of *rbcL* in this study did not interfere with the large stem loop, but the small, second, stem loop. The new secondary structure introduced with pRF+54c was, prior to the making of the chimeric gene construct, determined by using the RNA- and DNA folding software program Quikfold, developed by Dr. Michael Zuker. This showed that a stem loop with a shape and size *similar* to the original second small stem loop formed (Figure 13).

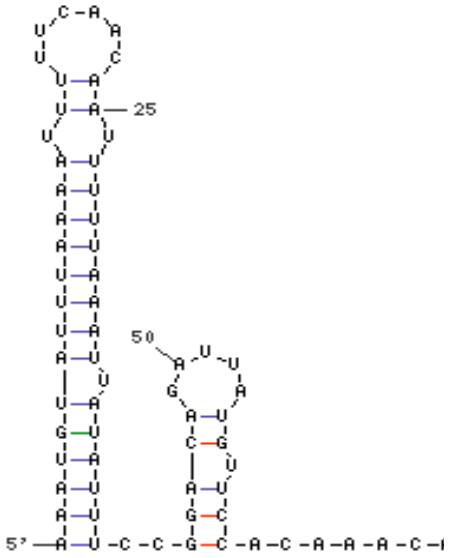

Figure 13: Secondary structure of 5´ UTR of the final modified plasmid pRF+54c as determined by the software program Quikfold. When compared to Figure 3 one can see that the conformation of the last four nucleotides (GGAC) of the endogenous rbcL 5' UTR structure, previously believed to be crucial for transcript stability, is disturbed. Other than that, the new second stem loop is quite similar, both in shape and size, to the original one.

3.2.2 Oligonucleotides

Based on the results of the previous studies described in section 1.1 in the introduction, and on the prediction of secondary structures by the mfold program, a decision was made to make a deletion of all bases between position +54 and +95 (not including +54 but including +95), relative to the TSS (+1). Thus, oligo-nucleotides composed of the original *rbcL* sequence from position +41 to +54 connected to the sequence from position +96 to +157, also from the *rbcL* gene, were ordered. The 5' to 3' strand of the oligonucleotide was designed to have a *BspEI* overhang (T´CCGGA) in its 5' end and a *PstI* overhang (CTGCA´G) in its 3' end (Figure 14). Both these overhangs were essential for completion of the first cloning step where +54 was cloned into SK+/ MU21 by digestion with these two enzymes. Oligonucleotides were ordered from MWG-Biotech AG.

```
5´CCGGACAGATTATGTTCCACAAACAGAAACTAAAGCAGGTGCTGGATTCAAAGCCGGTGTAAAAGACTACCGTGCA 3´
    3´TGTCTAATACAAGGTGTTTGTCTTTGATTTCGTCCACGACCTAAGTTTCGGCCACATTTTCTGATGGC 5´
```

Figure 14: Oligo nucleotide +54. Note the BspEI (T´CCGGA) and PstI (CTGCA´G) overhangs, on the left and right ends respectively, of the 5´- 3´strand.

500 pmol (5 µl à 100 pmol/µl) of each of the individual complementary and single stranded oligonucleotides (5` strand and 3` strand) were, according to the protocol by Sambrook and Russell (2001), phophorylated by T4 polynucleotide kinase, and run on a 1.3 % agarose gel. The oligonucleotides, from now on referred to as +54, were then isolated from the gel. This was done by cutting a small well in the gel in front of the +54 band, approximately the same size as the band, and placing a piece of dialysis tubing down in the well to stop the DNA from migrating further. The gel was then partly submerged in electrophoresis buffer while the cut out well was filled with buffer to ensure good conductivity. By running the electrophoresis a little longer, the DNA concentrated in the well and could be removed by pipetting. The DNA was then purified and precipitated by standard phenol/chloroform extraction and ethanol precipitation.

Phenol/chloroform extraction is an easy way to remove proteins from nucleic acid samples and can be carried out in a manner that is very close to quantitative. Nucleic acids remain in the aqueous phase and proteins separate into the organic phase or lie at the phase interface.

Normally this is done in Eppendorf tubes, in volumes between 100 and 700 µl. Here phenol-chloroform was added in a volume equal to the volume of each sample and vortexed until the solution got a white color.
The sample was then centrifuged at 13000 rpm for 2 minutes at room temperature, and the supernatant transferred to a new tube by pipetting. The procedure was repeated, but with chloroform only. Once the proteins were removed from the solution, the nucleic acids could be precipitated from the solution. DNA was recovered from the aqueous solution by addition of Na-acetate to a final concentration of 0.3 M (or a 0.1 final volume). Other types of salt can also be used for this purpose, but often at different concentrations. An appropriate volume of 96 % EtOH was added to the sample before incubation at - 80°C for approximately 1 hour, followed by centrifugation at 12000 rpm for five minutes at - 4°C. To get rid of all the salts in the solution, the sample was washed twice with 70 % EtOH. The DNA sample was re-suspended in 15 µl dH_2O, before storage at -20°C until needed.

Concentration was measured by the dot spot method. The dot spot method is a simple method for quantification of nucleic acids. 2 µl of the DNA sample is diluted 20-, 40-, and 80-fold in dH2O and spotted onto a UV-transparent plate (*e.g.* a Petri dish). Equal volumes of a series of DNA concentration standards (2.5, 5, 7.5, 10, 15 and 20 µg/ml) are spotted in an ordered manner on the same plate. To all spots 2 µl of a solution made from TE buffer and Ethidium bromide (EtBr) (2 µg/ml) are added and mixed by agitating with a pipette. A blank control consisting of 2 µl dH2O and 2 µl TE-EtBr solution is used for comparison. All spots are then observed under UV-light and since EtBr binds to DNA and take on a pink color under UV-light, the color intensity of the sample spots can be compared with the color intensity of the DNA standards (Sambrook and Russell 2001).

3.2.3 Isolation, digestion and ligation of the starting plasmid SK+/ MU21

The first cloning step in the construction of the final plasmid pRF+54c, involves the ~5 kb large vector SK+/ MU21. It is constructed from the original pBluescript SK+ vector and the MU21 vector by exchanging the original *XhoI-XbaI* fragment in pBluescript SK+ with the *XhoI-XbaI* fragment from MU21 (Figure 15).

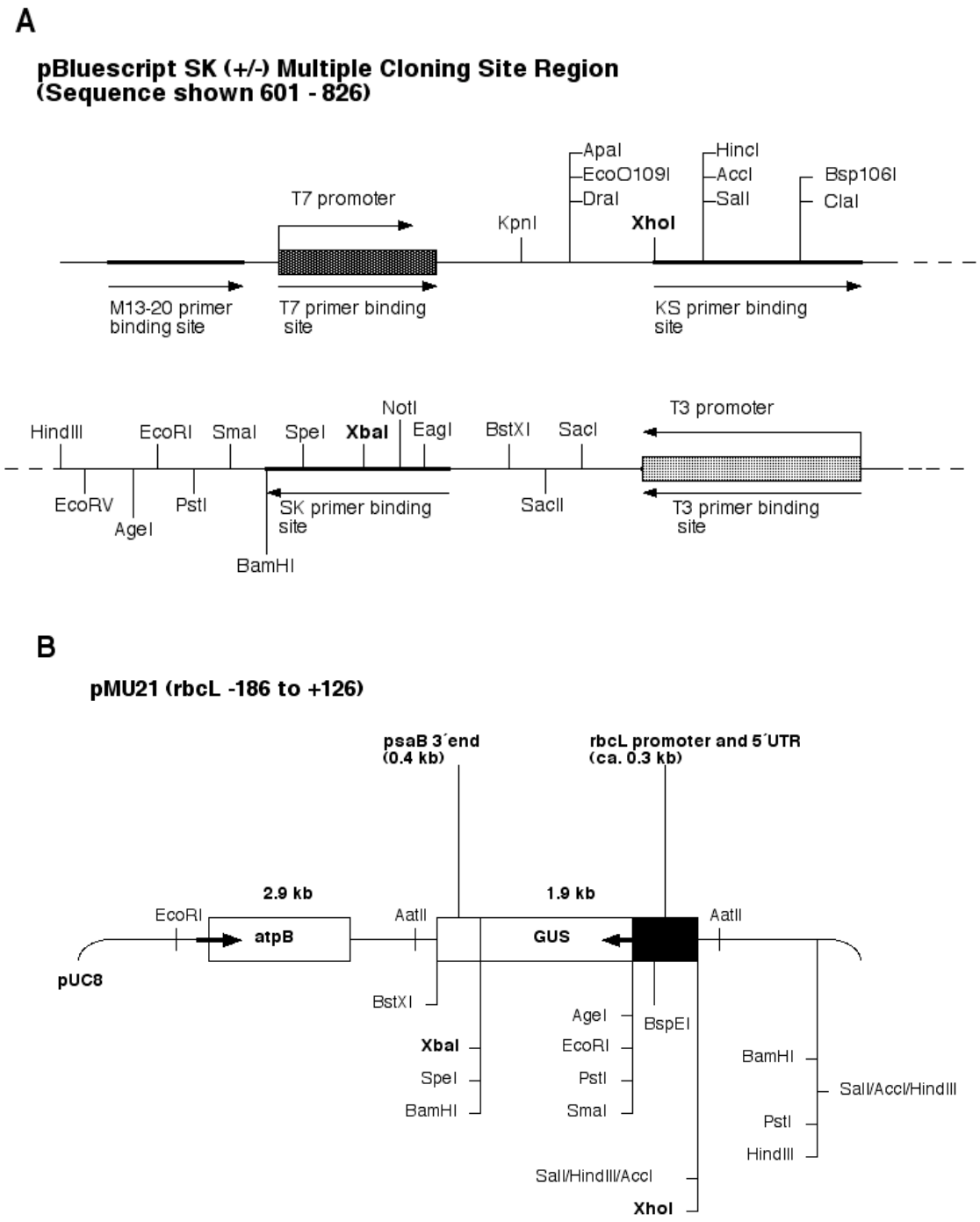

*Figure 15: Making of the SK+/MU21 vector. Both pBluescript SK+ (A) and p*MU21 *(B) were cut with XhoI and XbaI. The XhoI-XbaI fragment from p*MU21 *then replaced the original XhoI-XbaI fragment in pBluescript SK. Modified from http://www.fermentas.com (see section 8 for full web address) and from a figure of pMU21 provided by Uwe Klein. Figure is not drawn to scale.*

SK+/ MU21 was digested with *PstI* and *BspEI* for two hours in a 37 °C water bath and ran on a 1 % agarose gel. The larger band was then isolated from the gel in accordance with the procedure described in the text above, and its concentration measured by the dot spot method (see section 3.2.2 for details).

After isolation, 16 ng of the +54 oligonucleotide and 200 ng of the large fragment from the digested SK+/ MU21 were ligated together by T4-DNA ligation (Figure 16). When ligating by T4 DNA ligation the vector and insert were, in an Eppendorf tube, brought to a total volume of 6.5 µl with dH_2O. The tube was then incubated in a 45 °C water bath for 5 min. to avoid self-annealing by H-bonds, and immediately placed on ice. To the tube, 1 µl of thoroughly thawed and dissolved T4-DNA ligase buffer, together with 2 µl PEG (in a 30 % solution) was added. This brought the total volume to 9.5 µl. To this, 0.5 µl T4-DNA ligase was added. The tube was then incubated at a temperature between 16 °C (ideally) and room temperature for approximately 3 hours. The new plasmid was named pRF+54.

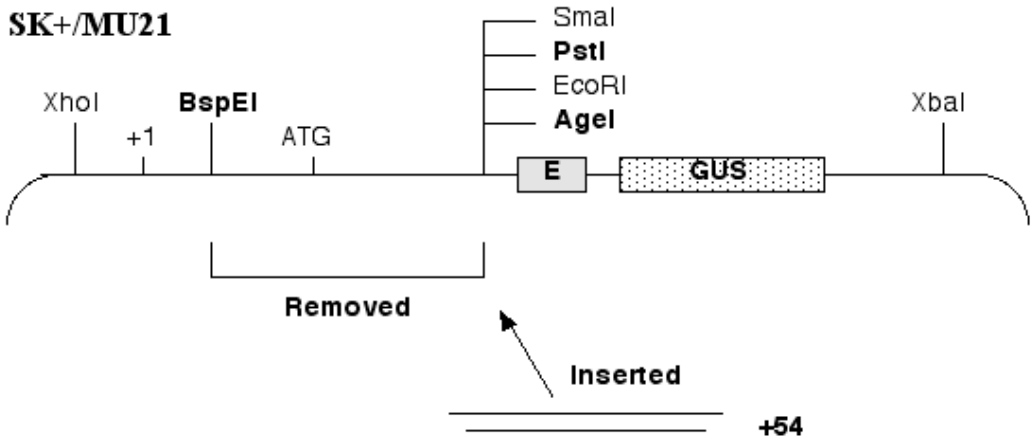

Figure 16: First cloning step and construction of pRF+54. Oligo nucleotide +54 was ligated into SK+/ MU21*, between the restriction sites BspEI and PstI by T4 DNA ligase, replacing the original region. E represents the enhancer region, while +1 indicates the transcription start site (TSS). Figure is not drawn to scale.*

3.2.4 Transformation of E. coli and control of plasmid

An aliquot of frozen competent *E. coli* cells in an Eppendorf tube, prepared by Uwe Klein according to the $CaCl_2$ procedure described in Sambrook and Russell (2001), was melted on ice, and 3 µl of ligation mix was added to it. Following the heat shock method for transformation (Sambrook and Russell 2001), which temporarily allows for incorporation of foreign DNA into the *E. coli* genome, the ligation mix and the competent cells were left on ice for 30 minutes before being exposed to 45 °C for 90 seconds and then re-immersed in ice. 800 µl of sterile LB medium was added to the tube, which was incubated at 37 °C for one hour. 75 µl of the mix was then spread on each of six LB + A plates, all of which were incubated at 37 °C over night. Since the SK+/ MU21 vector contains an ampicillin resistance gene, only transformed cells could grow.

Transformants were selected and grown over night in liquid LB + A medium at 37 °C. The plasmid was then isolated and amplified according to the miniprep protocol (Sambrook and Russell 2001). To verify that the +54 oligo was indeed inserted, both the transformed plasmid and the unmodified SK+/ MU21 vector were digested with *AgeI*. An *AgeI* restriction site is located immediately upstream of the *PstI* site in SK+/MU21. If the original region between *BspEI* and *PstI* in the vector was successfully replaced by +54, this restriction site would be removed and the newly constructed plasmid should remain uncut. The unmodified SK+/ MU21 vector, which still has the *AgeI* site was expected to be cut once and appear on the gel as one ~5 kb fragment, while the expected fragment for the uncut modified plasmid should be slightly smaller, but only by 41 bp, which is almost impossible to distinguish between on a gel when the fragments are this large. Though, since the uncut plasmid still will retain its circular form, the two forms will be able to be identified.

Once +54 was confirmed inserted, the plasmid was renamed pRF+54 and amplified again, this time by the maxiprep protocol for *E. coli* plasmid preparation by CsCl density gradient centrifugation (Tanaka 1975; Sambrook and Russell 2001). 10 µl from the miniprep culture from the selected transformants was added to 100 ml LB + A, and grown on a shaker at 37 °C over night. Procedure was according to protocol (Sambrook and Russell 2001). EtBr was removed from the plasmid preparation by repeated extraction with isopropanol followed by dialysis against TE-buffer. After preparing a sample by adding 10 µl from the DNA solution to 990 µl distilled water in a quartz cuvette, DNA was quantified by

spectrophotometry. At this point, 2 µg plasmid was sequenced as a final verification that the insert was indeed in.

3.3 Construction of plasmid pRF+54c

3.3.1 Isolation, digestion and control of pRF+54

A transformation vector, 32 / +10, and pRF+54 were both digested with *XhoI* and *XbaI* at 37 °C for two hours. *XhoI* and *XbaI* were used to liberate the entire *rbcL*:GUS cassette from pRF+54. Once run on a 1 % agarose gel, the large fragment from 32/+10 and the smallest fragment from pRF+54 were isolated from the gel, then purified and precipitated by standard phenol/chloroform extraction and ethanol precipitation. Concentration was measured by the dot spot method (Sambrook and Russell 2001) (see section 3.2.2 for details).
800 ng vector (the large fragment from 32/+10) and 300 ng insert (*rbcL*:GUS cassette) were ligated together by T4 DNA ligase (Figure 17) (see section 3.2.3 for details).

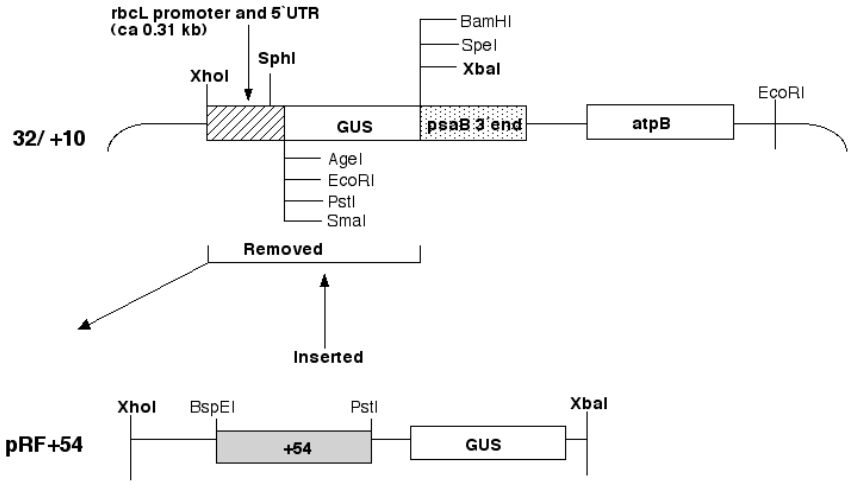

Figure 17: Second cloning step and construction of pRF+54c. The entire rbcL:GUS cassette is liberated from 32/+10 by digestion with XhoI and XbaI, and replaced by pRF+54. The figure is not drawn to scale.

3.3.2 Transformation of E. coli and control of plasmid

Following the same protocol procedure as described in section 3.2.4, pRF+54c was transformed into *E. coli* by the heat shock method. 0.8 ml sterile LB + A medium was added to the transformed cells, and 75 μl was spread onto each of four LB + A plates and incubated at 37 °C over night.
Transformed cells were selected and grown in 3 ml liquid LB + A, and plasmid was isolated from the cells according to the miniprep protocol (Sambrook and Russell 2001).

The 32/+10 vector has an *SphI* site around its position +10, and if pRF+54 was indeed inserted, this site should be removed since it would replace this area in 32/+10.
To confirm the insert, the transformed plasmid and unmodified 32/+10 were digested with *SphI* and run on a 1% agarose gel. If the insert was in, pRF+54c should remain uncut while 32/+10 should be cut once (see section 4.1.2 for details).

Once confirmed, the newly constructed plasmid was named pRF+54c.

The final chimeric gene construct, pRF+54c, could then be transformed into the *Chlamydomonas* chloroplast. Two successfully transformed plasmid samples were chosen and amplified by maxiprep according to the protocol described at the bottom of section 3.2.4. DNA was quantified by the dot spot method (see section 3.2.2 for details). As a final verification of the construct the plasmid was sequenced.

10 μg of DNA from the maxiprep of each of the two chosen transformants was precipitated for chloroplast transformation.

3.4 Chloroplast transformation of Chlamydomonas

Mutants were grown in HSHA, spun down and re-suspended in 500 μl HSHA. 100 μl re-suspended cells were transferred to 1 ml low-density HSHA agar holding at 40 °C, mixed, and quickly transferred to HSHA agar plates. Plates were kept in the dark for five hours prior to transformation. pRF+54c was introduced into the chloroplast by particle bombardment (Figure 18), according to the procedure described by Boynton and Blowers (Blowers *et al.* 1989; Boynton *et al.* 1988).

Figure 18: The PDS-1000 He particle bombardment system from BioRad used to transform the Chlamydomonas chloroplast. From www.fisio.dipbsf.html (see section 8 for full web address).

CC-373 cells were bombarded with DNA precipitated onto 0.6 µm gold particles at high velocity by a burst of helium gas at a pressure slightly higher than 1350 psi, from a distance of 15 cm.

Once bombarded, the plates were kept in the dark over night to allow for the transformation vector pRF+54c to be integrated into the chloroplast genome by homologous recombination, completing the *atpB* gene and the IR. Subsequently, the cells were transferred to HS plates and gradually exposed to bright light for growth.

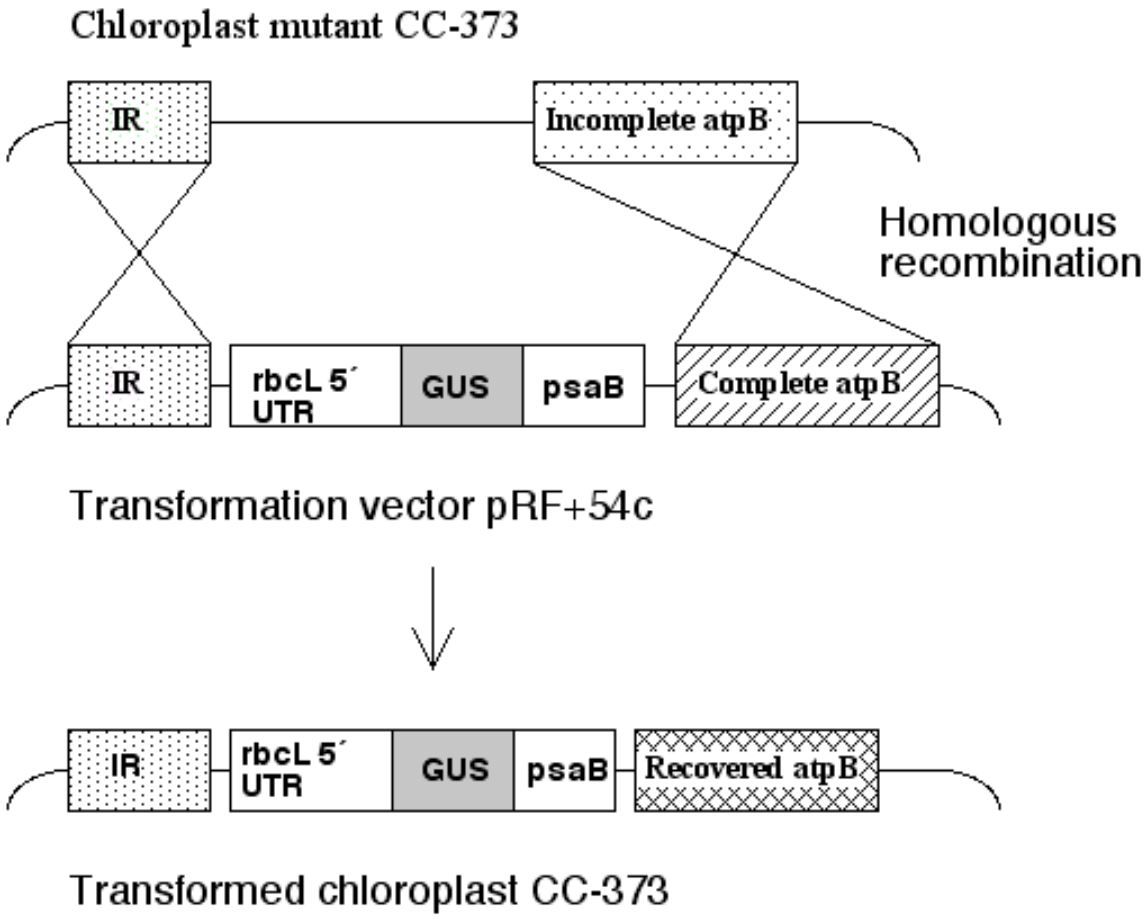

Figure 19: Integration of pRF+54c into the CC-373 chloroplast genome. Homologous recombination between the incomplete atpB gene and IR in CC-373 and the complete atpB and IR in pRF+54c restores the photosynthetic abilities of CC-373.

The completion of the *atpB* gene restores the photosynthetic capacity of the CC-373 cells, enabling selection of transformants since only photoautotrophic cells can grow on HS medium (Figure 19). The transformed cells can be seen as small and green colony forming units (CFUs) amongst the white background of dead, untransformed cells. As many transformants as possible were transferred to a fresh HS plate and maintained under light over time. Unfortunately, verification of successful transformants cannot solely be based on this. This is due to the fact that recombination between vector and the chloroplast genome of CC-373 can restore both the *atpB* gene and the IR without actual

integration of the reporter gene construct. Therefore it is necessary to quantify the GUS content of each transformant.

3.5 Slot blot analysis – A qualitative determination of GUS content

To determine the relative GUS content in the selected transformants, slot blot analyses were carried out. Seven transformed CFUs were inoculated in 100 ml HS each and grown in Erlenmeyer flasks, under light, for four days. Approximately 75 ml of each culture was transferred to large glass tubes, supplemented with fresh liquid HS and grown for four days in a 32 ºC water bath, constantly supplied with air and 2% CO_2 (
Figure 20)

Figure 20: Growth of transformed CC-373, in HS medium in a 32 ºC water bath, for later isolation of genomic DNA and RNA. Cells were constantly supplied with air and 2% CO_2.

Following the protocol (Dellaporta *et al.* 1983), genomic DNA was isolated from the selected transformants. Concentration of the isolated DNA was determined by the dot spot method (see section 3.2.2 for details) (Sambrook and Russell 2001).

500 ng DNA from each transformant was in an Eppendorf tube, together with 3 μl 5N NaOH, brought to a 50 μl volume by addition of dH_2O. All tubes were incubated at 65 ºC for 45 minutes for denaturing, then cooled down to room temperature before 50 μl 20X SSC buffer (sodium chloride-sodium citrate in distilled/deionized water) was added to each tube. The Hoefer slot blot apparatus was assembled and the sample DNA blotted onto a nylon membrane (ZetaProbe; BioRad). Once the DNA is blotted, the nylon membrane is wrapped in cling wrap and cross-linked under UV light, DNA side facing the light, for two minutes. This is to bind the DNA to the membrane. The membrane was then probed with a radio labeled GUS probe generated by the random primer labeling method (Feinberg and Vogelstein 1983). The membrane is equilibrated in pre-hybridization buffer before it is immersed in hybridization buffer and hybridized over night, or longer.

The membrane is then exposed to X-ray film at -80 ºC in a light-impermeable metal cassette. Exposure time can vary, depending on both the activity of the probe and on the abundance of target for the probe, but generally 24 hours is sufficient. Naturally, the same factors affect the signal strength. Because of this, slot blots only give a qualitative description of the GUS content in the DNA from the transformants. Still, by comparing the signal strengths, a relative determination of GUS content in each sample can be made.

3.6 Southern blot analysis – A quantitative determination of GUS content

After analyzing the relative GUS contents of the samples, the transformants with the highest relative GUS content were chosen. Again, genomic DNA was isolated as described in section 3.5. As mentioned before, restoration of photosynthetic abilities by completion of the *atpB* gene can occur without integration of the reporter gene construct. This results in a restriction fragment length polymorphism when digesting with *KpnI* and *HindIII*. If the reporter gene construct is indeed integrated, the resulting restriction fragment will be ~5.4 kb, compared to ~3 kb if it is not.

To determine the degree of homoplasmicity with regard to the GUS sequence, a Southern blot was performed. There are various procedures for Southern blots, but all are anchored in the same principles; digestion of DNA, electrophoretic separation by size on a gel, and transfer of DNA to a membrane, usually by capillary pulls.

In this study, 2.5 μg of each sample was digested with *KpnI* and *HindIII*, and run on a 1 % agarose gel. In accordance with the protocol for alkaline blotting (described in the BioRad ZetaProbe manual) the DNA was transferred to a membrane (Zetaprobe; BioRad) and probed with a radio labeled *atpB* probe. The procedures for probing, hybridization and washing were essentially the same as they were for Slot blot.

3.7 Northern blot analysis - Reporter gene transcript accumulation

Northern blot is similar to Southern blot analysis except the fact that in Northern blots RNA, rather than DNA, is transferred to the membrane. Therefore it is necessary to keep all solutions and equipment RNase-free since RNases are quite abundant enymes that break down RNA. Also, since RNA has tendencies to form secondary structures it is necessary to apply denaturing conditions.

Northern blot is used to detect the reporter gene transcript and to see if it accumulates in the cell. Total cellular RNA was isolated according to protocol (Sambrook and Russell 2001) after 11 hours in the dark in a 12 hour light / 12 hour dark cycle. Absorbance was measured at 260 nm by spectrophotometry and concentration of the sample was calculated (OD 1.0 = 40 µg RNA/ml). 4 µg of RNA was run on a 1.3 % denaturing agarose/formaldehyde gel, essentially as described by Blowers *et al*. (1989).

RNA was transferred from the gel to a nylon membrane (ZetaProbe; BioRad) according to the protocol provided by the manufacturer, and probed with a radioactive GUS probe before 24 hours of exposure to an X-ray film at -80 ° C. As controls, MU7 and the unmodified +157, carrying the original *rbcL* 5´ UTR, was used. Also here the procedures for probing, hybridization and washing were essentially the same as they were for Slot blot and Southern blot.

4 RESULTS

4.1 Cloning of pRF+54c into the Chlamydomonas chloroplast

4.1.1 Construction of pRF+54

Once the secondary structure of the modified *rbcL* 5´ UTR was predicted (see Figure 13) and the two single stranded oligonucleotides were annealed and named +54, both +54 and SK+/MU21, which was digested with *PstI* and *BspEI*, were run on a 1.3 % agarose gel. +54 and the larger band from SK+/MU21 were isolated, essentially as described in section 3.2.2.

+54, which was created to have *PstI* and *BspEI* overhangs, was then easily ligated into the SK+/MU21 vector by T4 DNA ligation (see section 3.2.3 for details). The construct was transformed into *E. coli* (essentially as described in section 3.2.4) and 75 µl of transformed cells were spread onto each of eight LB+A plates, which gave close to 70 colonies. Eight of these were selected and amplified by miniprep.
The cloning of +54 into SK+/MU21 leads to the deletion of an *AgeI* restriction site around position +150 in SK+/MU21 (Figure 16). To verify that +54 was indeed inserted, a test cut with *AgeI* was performed on DNA from each of the selected transformants (Figure 21).

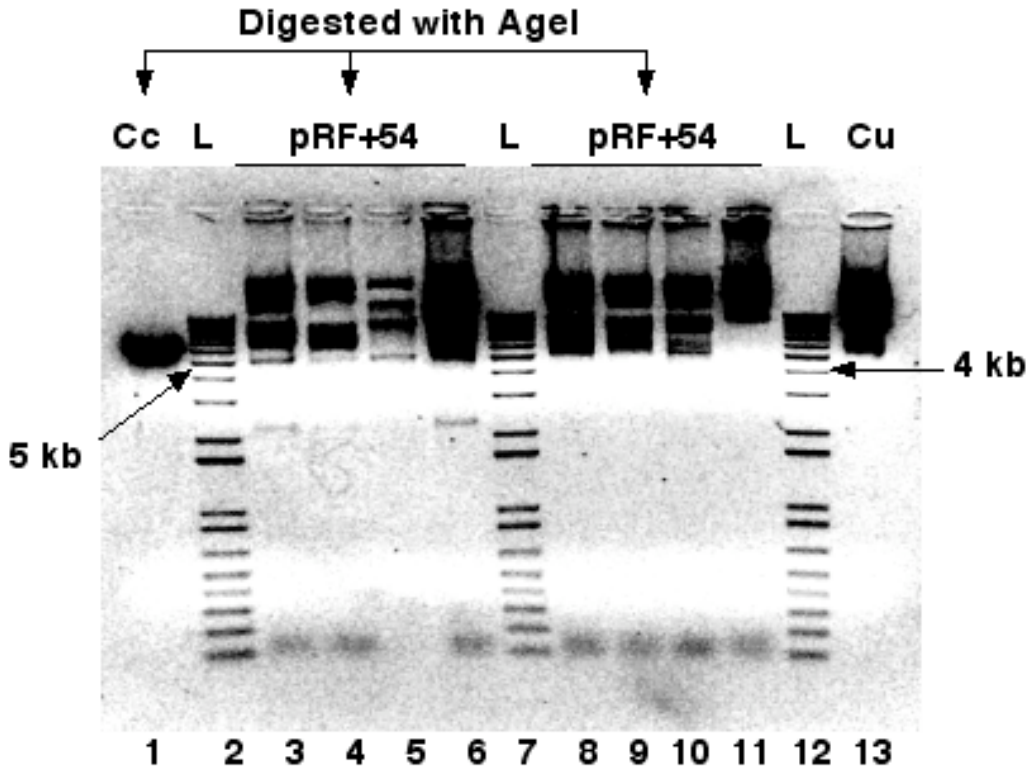

*Figure 21: Test cut of pRF+54 to verify that +54 was successfully cloned into SK+/*MU21. *Lanes 3-6 and 8-11 represent miniprep samples, all digested with AgeI. In lanes 2, 7 and 12 a 1 kb plus ladder (L) was run. Unmodified SK+/*MU21 *was run as a control, both digested and undigested, in lanes 1 and 13 respectively (annotated as Cc for cut control and Cu for uncut control). If +54 was inserted into SK+/* MU21 *the DNA should remain uncut since the AgeI restriction site is removed with the insertion of +54. It is clear from the picture that minipreps 1 through 4 and 8 through 10 are uncut like the control in lane 13, ergo +54 is inserted. The miniprep in lane 11 seems to be very large, around 11 kb, and cannot be the same plasmid as the rest. This cannot be used. The arrows on the sides point to the 4- and 5 kb fragments of the 1 kb plus ladder. These show that the restriction fragments are of the correct, expected, sizes, which is around 5 kb. The arrows on the top indicate which samples were digested with AgeI.*

Once the selected transformants were screened for the +54 insert, the sample in lane 9 was chosen to proceed with, and thus amplified by maxiprep. 2 μg of this selected cloned plasmid were sent for sequencing and the sequencing results gave a final confirmation that

+54 was successfully cloned into SK+/MU21 (Figure 22). This new plasmid, constructed from SK+/MU21 and +54 was named pRF+54.

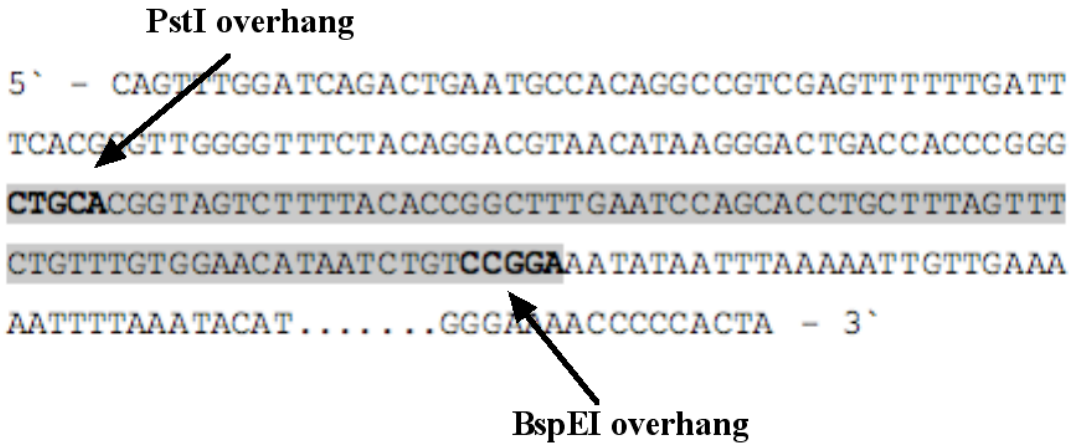

Figure 22: The figure shows the sequencing results of the 5`- 3` strand from pRF+54. The +54 oligonucleotide is highlighted in grey, and the two arrows point to the constructed PstI and BspEI overhangs (CTGCA and CCGGA respectively).

As described in section 3.2.2, insertion of +54 into the *rbcL* gene leads to a deletion of the region between positions +54 and +95 (not including +54 but including +95) in the original *rbcL* sequence (Figure 23 and figure 24).

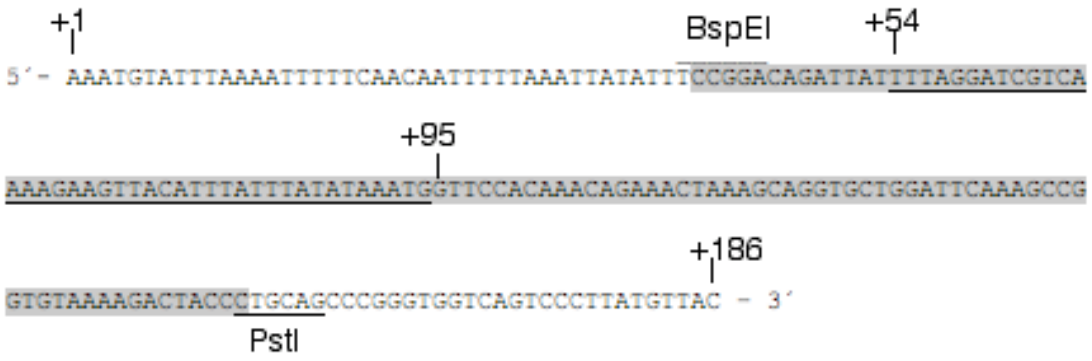

*Figure 23: The figure shows the original rbcL sequence in SK+/*MU21 *from position +1 (TSS) to position +186. Highlighted in gray is the region that is cut out when digesting with BspEI and PstI, while the underlined sequence is the region being deleted (from +54 to +95).*

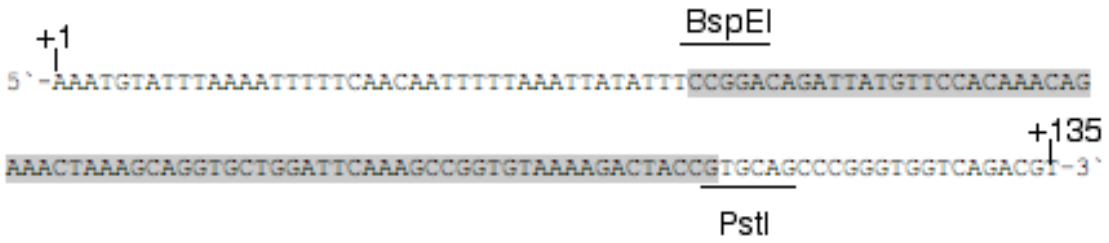

Figure 24: The figure shows the new sequence of the rbcL 5´ UTR in SK+/MU21 after the deletion, once +54 is cloned in.

4.1.2 Construction of pRF+54c

pRF+54 and transformation vector 32/+10 were both digested with *XhoI* and *XbaI (*Figure 25*)*. The fragments were run on a 1.3 % agarose gel and isolated (see section 3.2.2 for procedure).

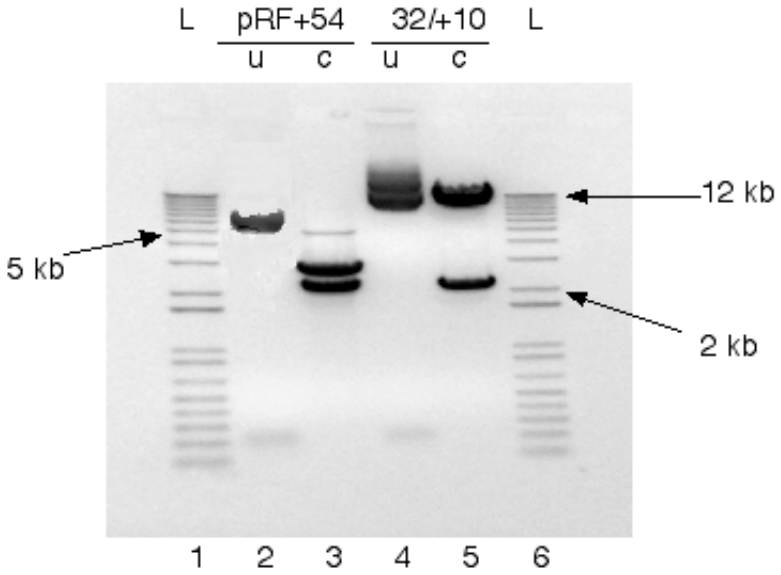

Figure 25: Digestion of pRF+54 and 32/+10 with XhoI and XbaI. The picture shows pRF+54 and transformation vector 32/+10 as undigested (lanes 2 and 4) and as digested with XhoI and XbaI (lanes 3 and 5). In lanes 1 and 6 a 1 kb plus ladder (L) was run. The smallest (bottom) band from pRF+54 in lane 3 and the largest (top) band from 32/+10 in lane 5 were both isolated from the gel and ligated together by T4 DNA ligase, as described in detail in section 3.2.2. The numbers 1 through 6 at the bottom of the pictures indicate the lane numbers. Arrows point to the 2-, 5- and 12 kb fragments of the ladders.

As described in detail in section 3.3.1, pRF+54 was ligated into the transformation vector 32/+10 and the resulting new construct, named plasmid pRF+54c, was transformed into *E. coli*. Again, 75 µl of transformed cells were spread onto each of eight LB+A plates resulting in 28 colonies of transformants. Plasmid DNA was eventually isolated from six of these colonies and amplified by miniprep.

32/+10 contains an *SphI* restriction site at position +10, hence the name, and, if pRF+54 was successfully inserted, this site should be removed. To check this, a test cut with *SphI* was performed on the six miniprep samples (Figure 26).

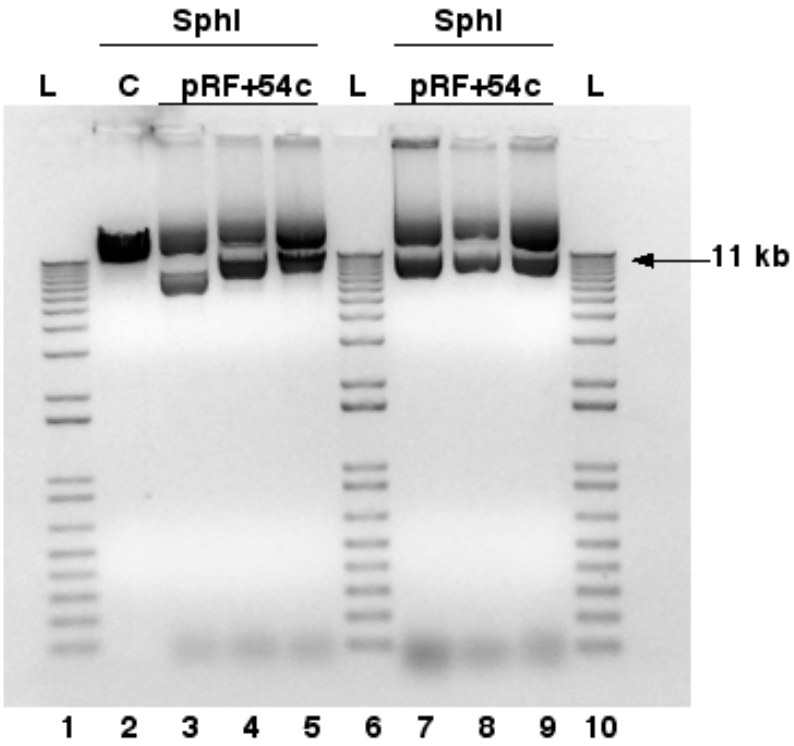

Figure 26: Test cut of pRF+54c to confirm insertion of pRF+54 in transformation vector 32/+10. Lane numbers are indicated at the bottom of the picture (1-10). Lanes 1, 6 and 10 show 1kb plus ladders (L). Lane 2, annotated C, shows the 32/+10 vector digested with SphI, while lanes 3-5 and 7-9 show DNA from miniprep samples (see section 3.3.2 for details) also digested with SphI. From analyzing this picture one can see that the control in lane 2 is cut, while all samples are uncut. The lower band in lane 3 has migrated farther than the others for unknown reasons, perhaps due to very supercoiled form. Proceeded with samples in lanes 4 and 5.

After confirming the composition of the final plasmid, pRF+54c could be transformed into the *Chlamydomonas* chloroplast by micro-projectile bombardment.

As described in section 3.4, the bombarded cells were kept in the dark over night to allow integration of pRF+54c into the chloroplast chromosome before they were transferred to HS plates for growth. 26 transformants grew, appearing as green colonies on a background of dead white cells, and all were eventually transferred to the same HS plate and assigned the numbers 1-26. Genomic DNA was isolated from seven selected transformants (number 3, 5, 6, 13, 18, 21 and 26) and the concentration of the DNA from each clone was measured by the dot spot method (see section 3.2.2).

4.2 Slot Blot – A qualitative determination of GUS content

DNA from the selected transformants was loaded onto the slot blot apparatus in a specific and recognizable pattern, for easy orientation of the membrane. Once the DNA was blotted onto a nylon membrane it was probed with a radiolabeled GUS probe that will bind to a certain sequence in the GUS gene, and exposed to X-ray film at – 80 °C for approximately 48 hours. After hybridization and washing the result of the blot could be analyzed. The intensity of each band is an indication of the GUS content (Figure 27).

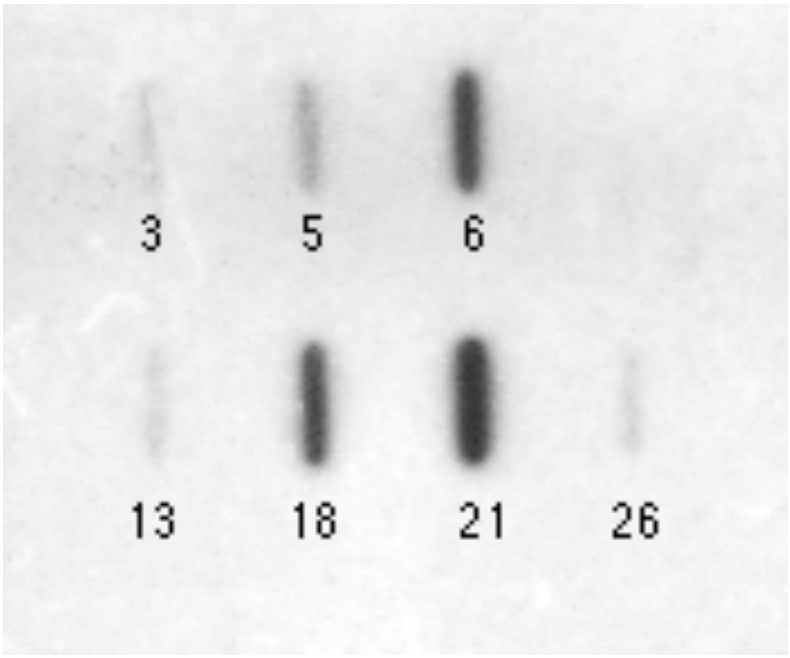

Figure 27: Picture of autoradiogram from Slot Blot. The samples were loaded as described in the text above and the numbers indicate which clone the blotted DNA is from. The intensity of each band is an indication of the relative GUS content of that particular clone; high intensity indicates high GUS content. All transformants had the GUS gene but in varying amounts. Clone number 21 had the highest content followed by number 6 and 18. The GUS content in the remaining clones, number 3, 5, 13 and 26, were too low to proceed with.

All seven clones harbored the GUS gene but in varying amounts. Clone number 21 had the highest content followed by number 6 and 18 respectively, and these three clones were chosen to proceed with. The remaining clones` GUS contents were too low to proceed with.

4.3 Southern blot – A quantitative determination of GUS content

As mentioned in sections 3.4 and 3.6, transformation of the *C. reinhardtii* chloroplast can render two forms of genomes, both leading to completion of the *atpB* gene and restoration of photosynthetic capabilities in *C. reinhardtii (*Figure 28*)*. One form will have the chimeric

*rbcL:*GUS reporter gene integrated while the other form will not. The pool of copies of the chloroplast genome in one cell, usually 50-80 copies, will consist of a mixture of these two forms, and the ratio between the two in cells from the same transformant will vary. Digestion with *KpnI* and *HindIII* will make it possible to separate these to forms from one another since *KpnI* cuts in the IR located upstream of the *rbcL* promoter and *HindIII* cuts right upstream of the *atpB* promoter. If the chimeric *rbcL:*GUS gene construct is integrated this will result in a ~5.4 kb fragment while a ~3 kb fragment will be the result if it is not integrated.

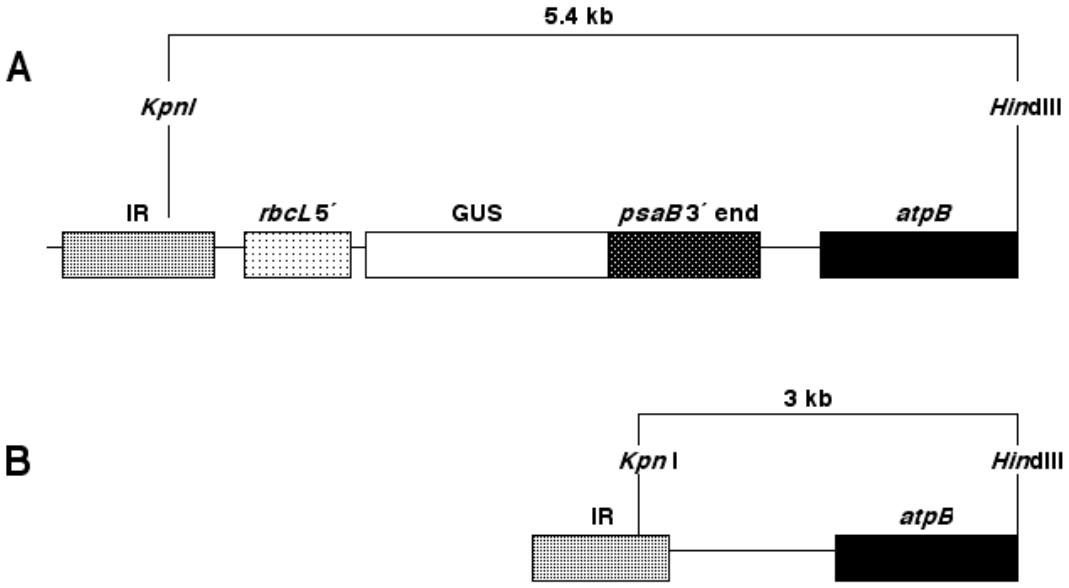

Figure 28: The two forms of genomes that can exist in the Chlamydomonas reinhardtii chloroplast after micro projectile bombardment of the CC-373 with pRF+54c. A: the rbcL:GUS chimera is integrated, resulting in a ~5.4 kb fragment when digested with KpnI and HindIII. B: The rbcL:GUS chimera is not integrated, resulting in a ~3 kb fragment when digested with the same enzymes. Both forms have the completed atpB gene, leading to restoration of photosynthesis.

Since both fragment types will contain the completed *atpB* gene, a radio labeled *atpB* probe was used was used to measure the ratio between the two. This probe is a 0.7 kb

HpaI/*EcoRV* restriction fragment released from a vector, pCrc*atpB*, which contains an internal portion of the *Chlamydomonas reinhardtii* chloroplast gene *atpB*. Based on the intensities of the bands from the two types of fragments in the autoradiograph, the ratio between the two forms of chloroplast genomes can be determined.

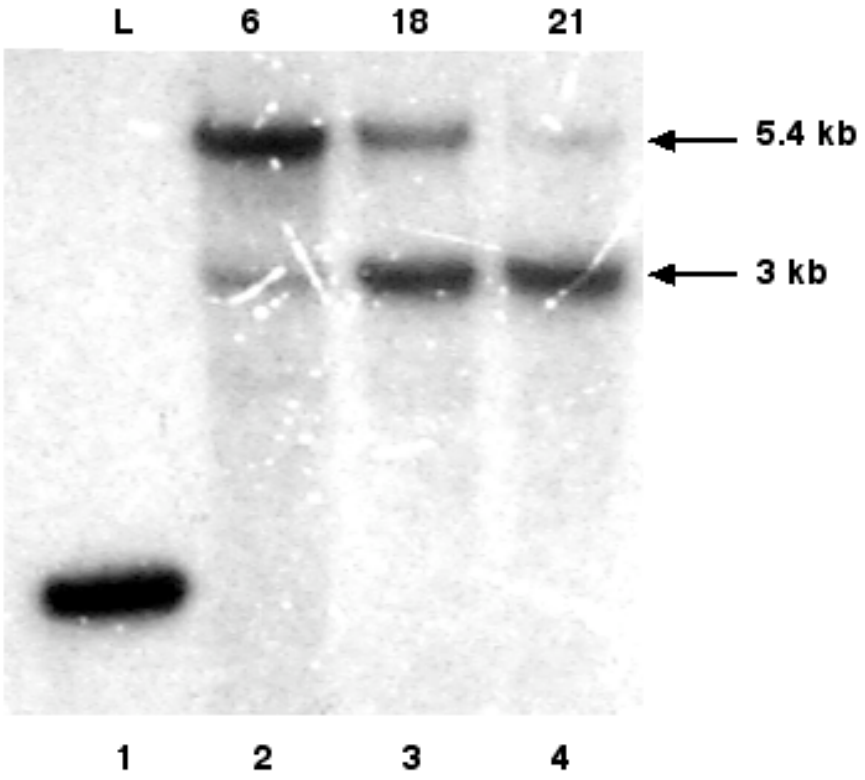

Figure 29: Autoradiogram from Southern blot. DNA isolated from the three selected transformants (clones 6, 18 and 21) was digested with KpnI and HindIII producing fragments of two sizes, depending on whether the rbcL;GUS chimera was successfully integrated into the Chlamydomonas chloroplast or not. The numbers on top of the figure indicates which clone the DNA came from while the numbers on the bottom indicates lanes 1-4 on the original gel. Clone 6 has a very intense ~5.4 kb fragment and a quite weak ~3 kb fragment, indicating a homoplasmicity of approximately 80 % in favor of the inserted rbcL;GUS chimera. Clone 18 has a homoplasmicity of maybe 40 % while clone 21has one less than 10 %. In lane 1 a 1 kb plus ladder was run as a control. The band in that lane has a size of ~1.6 kb and hybridizes very strongly to the atpB probe. Therefore it can be used as a hybridization control.

Southern blot was assembled and performed as described in section 3.6, and the membrane was exposed to an X-ray film at – 80 °C for 48 hours before the autoradiogram was developed (Figure 29). All clones tested (clones number 6, 18 and 21) contained both forms of the chloroplast genome (Figure 28) but in different ratios. Clone number 6 had the highest I:NI ratio (I representing integration of the chimeric *rbcL*;GUS reporter gene into the *C. reinhardtii* chloroplast, and NI representing no integration) of approximately 80:20. Clones number 18 and 21 had approximate ratios of 30:70 and 10:90 respectively. Although a homoplasmic clone with respect to "I" (100% integration) is preferable, clone number 6 has a high enough degree of homoplasmicity to perform Northern blot on. But one has to take a degree of homoplasmicity lower than 100 % into account when analyzing the result from the Northern blot.

4.4 Northern blot - Reporter gene transcript accumulation

To analyze the stability of GUS transcripts harboring the +54 deletion, accumulation of GUS RNA was determined by Northern gel blot. Accumulation of GUS transcripts in pRF+54c was compared to two controls, MU7 and +157, a Northern blot was carried out. MU7 is a chimeric reporter gene containing both the GUS gene and the unmodified *rbcL* 5´ UTR, but lacking the transcription enhancer sequence (see Figure 12). +157 is another chimeric reporter gene that harbors the GUS gene fused to the original unmodified *rbcL* 5´ UTR, and harbors, like pRF+54c, the transcription enhancer sequence, located around position +140. The inclusion of the enhancer region will significantly increase the transcription rate, so that factor will have to be taken into account when comparing the intensity of the bands. In a Northern blot there is a direct correlation between the intensity of the bands on the autoradiogram and the amount of target RNA transcripts accumulated. But, for this to be of any value when comparing, it is important that the added amount of RNA is the same for each sample. This was verified by comparing the intensity of the bands on the agarose gel under UV light before blotting, in addition to being accurate when measuring the concentration of, and preparing each sample. Samples were separated electrophoretically for approximately three hours at 60 mA. The bands that are visible on the gel are rRNAs since the amounts of mRNA in the samples are too small to be seen (Figure 30).

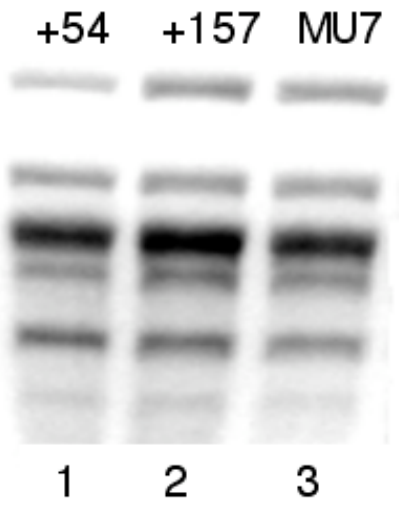

*Figure 30: Picture of the Northern agarose gel. **+54** indicates pRF+54c. The numbers on the bottom indicate lanes 1-3. One can see that the bands from +157 are a bit more intense than the others, suggesting a slightly higher DNA concentration. This must be taken into consideration when determining the levels of accumulation from the Northern blot.*

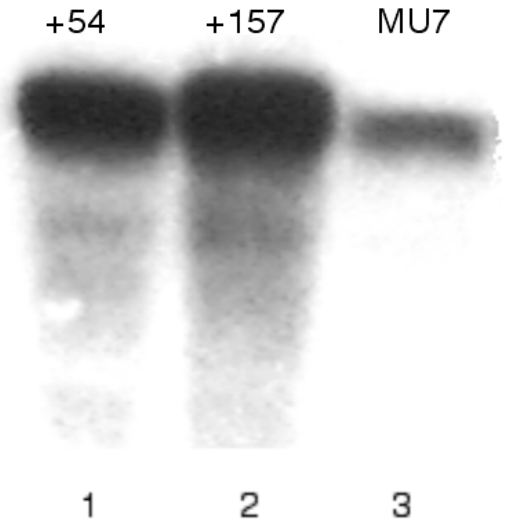

Figure 31: Autoradiogram of Northern blot. +54 indicates pRF+54c. GUS transcripts from +157 accumulate to the same levels as the endogenous rbcL gene in C. reinhardtii. . Based on the intensity of the bands one can see that GUS transcripts of pRF+54c accumulate almost to the same level as transcripts of SK+/157, while GUS transcripts from MU7 accumulates significantly less. The numbers on the bottom indicate lanes 1-3.

By comparing the intensities of the bands from pRF+54 and +157 in the autoradiogram (Figure 31) it becomes clear that GUS transcripts from pRF+54c accumulate to a very high, almost endogenous, level. To make a proper determination of the actual GUS accumulation, some things have to be taken into consideration. First, as described above, the concentration of +157 appears to be higher than that for MU7 and pRF+54c. When taking that, together with the fact that pRF+54c has a degree of homoplasmicity of 80 %, and that the intensity of the band from pRF+54c is approximately 80-90 % as intense as the band for +157, into account, it can be concluded that their transcripts accumulate essentially to the same levels. Transcripts from MU7 on the other hand accumulate to a significantly lower level due to it lacking the transcription enhancer sequence.

Based on these observations it can be concluded that pRF+54c does accumulate GUS transcripts. Thus, deletion of the nucleotides between +54 and +95, which leads to an alteration of the second smaller stem loop, does not affect transcript stability. The chimeric gene construct not only renders stable transcripts, but transcripts that appear to be *as stable* as those from the endogenous *rbcL* gene.

5 DISCUSSION

The results presented in this study show that a deletion of all nucleotides between position +54 and +95 (not including +54 but including +95) relative to the TSS (+1) in the *Chlamydomonas* chloroplast gene *rbcL,* does not affect transcript stability. In fact it appears that transcripts from the chimeric *rbcL;GUS* gene construct accumulate to similar levels as transcripts from the endogenous *rbcL* gene when integrated into the CC-373 genome.

As described in the introduction, this study is part of a larger project aiming to identify the sequence and conformation, and to narrow down the number of nucleotides, in the *rbcL* 5´ UTR required for transcript stability. Previously, whole portions of the sequence have been replaced, and nucleotides have been both added to and removed from the 5´ UTR of *rbcL* for this purpose (Suay *et al.* 2005; Salvador *et al.* 2004; Anthonisen *et al.* 2001) (see section 1.1 for details). As described in section 2.1 it has been found, contrary to what was previously believed, that in the 5' UTR of *rbcL,* the large stem loop itself does not confer stability on transcripts. Instead it mediates folding of a ten nucleotides long, single stranded region between +38 and +47 (relative to the TSS at position +1) around its base that has been found to be crucial for transcript stability (Suay *et al.* 2005). Previously it was believed that both the sequence and the conformation of this entire sequence were of importance. It has been proven that the *sequence* is, but only the conformation of the six first nucleotides of this region has been found important for transcript stability.

One study involved a 3' - 5' deletion, where nucleotides between positions +45 to +156 were deleted. This led to point mutations in positions +46 and +47, where an A was changed into a U, and a C was changed into a T, respectively. These mutations disturbed the sequence of this ten nucleotides long region, and rendered unstable transcripts, thus providing additional confirmation that this sequence is of importance for transcript stability. The secondary structure resulting from these modifications has not been determined. Therefore, although it is likely, it is unknown whether or not the modifications disturbed the structure of the sequence (Haugen P M 2004).

The deletion between +54 and +95 disturbs the conformation, but not the sequence, of the last four nucleotides (those in positions +44 to +47) of the ten nucleotides long region, but

still renders stable transcripts, so it is safe to say that the conformation of these last four nucleotides is not of significance for transcript stability.

This study also goes beyond the much mentioned 10 nucleotides long region and the large stem loop, and challenges the importance of another surrounding area *e.g.* the second, smaller, stem loop. Although not much focus has been aimed at the small loop, some speculations have been made as to whether or not this stem loop itself functions as a binding site for a trans-acting factor(s). In previous studies it has, as a contradictory element to this theory, been found that when the sequence downstream of +63, which is part of the small loop, was deleted, transcripts remained stable. Thus this sequence is not important for stability, nor is its structure. If we look at that, together with the results from *this* study where almost the entire endogenous second loop has been deleted, and then consider that the entire endogenous 5`UTR is a mere 92 nucleotides long, it is clear that more than half of the 5` UTR can be deleted without affecting transcript stability.

Since proteins that bind to the 5´ UTR may regulate processing or RNA stability, such binding activities have been sought and characterized. Often a link between translation and transcript stability has been found for various genes. As an example, a homologue of *the E. coli* ribosomal protein S1, consistent with a role in translation initiation, has been found to bind to the *psbA* 5´ UTR in vitro. It has been speculated that binding of this protein couples *psbA* processing and translation. In the *rbcL* 5`UTR a translation-important element, more specifically the ribosome binding site, is located in the deleted region. Since this is deleted without affecting transcript stability it can be concluded that for transcripts of the *rbcL* gene no such link exists.

5.1 Transcription

The strong, visible signals on the Northern blot autoradiogram, show that the pRF+54c construct renders stable transcripts. This obviously confirms that transcription did occur. But if the deletion rendered unstable transcripts, Northern blot would be a quite unreliable method for detecting transcription since it relies on accumulation of transcripts to produce signals on the autoradiogram. Thus it would be difficult to distinguish unstable transcripts from a total lack of transcription. In previous studies though, several changes have been made to the large stem loop, and nucleotides have also been added to the 5´ end of the 5´

UTR, without affecting transcript stability. Since a basic promoter element is essential for transcription, these experiments strongly suggest that no such element is located in this region. Thus it is safe to say that any changes in stability are due to alterations of sequence or structure, and not to changes in transcription.

+157 and MU7 are included as controls in the blot to show that the Northern blot procedure worked, and also to help estimate the accumulation of transcripts. +157 contains the endogenous transcription enhancer sequence and thus accumulates transcripts to endogenous levels, while MU7 does not contain the enhancer sequence and therefore accumulates less, only about 10 % of the accumulation levels of +157. This gives us two levels of accumulation for more accurate comparison and estimation.

5.2 Cis-acting elements and trans-acting factors

Some mRNAs have a half-life of only a few minutes, while others are long-lived with half-lives up to several hours. Therefore there must be factors involved, protecting the long-lived mRNAs from ribonucleolytic- or immediate degradation. For prokaryote-type mRNAs, important determinants for longevity have been delineated to primarily reside in the 5´ UTR (Suay *et al.* 2005). Therefore, it is natural to believe that trans-acting factors bind in that area.
Putative cis-acting elements have been identified for several genes located in the 5´ UTR of *Chlamydomonas,* spinach and other species, and the trans-acting factors binding to them have been isolated, often by genomic complementation (Boudreau *et al.* 2000). Another fact that supports the idea of putative cis-acting elements and trans-acting factors playing a role in mRNA stability is the existence of a large number of nuclear mutants of *Chlamydomonas* failing to accumulate specific individual mRNA populations in the chloroplast. These include chloroplast transcripts encoded by *atpA, atpB, petA, petB, petD, psbB, psbC* and *psbD* (Boudreau *et al.* 2000).

One reason why it is important to narrow down the number of nucleotides in the 5´ UTR of *rbcL* required for rendering stable transcripts is that it would minimize the area where trans-acting factors may possibly bind, making it easier to identify these. Many suggestions have been made as far as to where these factors may, or may not, bind. The fact that transcript stability depends on a specific sequence in a fairly specific conformation suggests that the

region between +38 and +47 is, or is part of, a cis-acting element, functioning as a binding site for a trans-acting factor(s) that protects transcripts from rapid degradation in chloroplasts.

The nucleotides in position +48 to +54 relative to the TSS (+1) are also of interest, so a natural next step would be to make a construct similar to pRF+54c where these bases are deleted as well. If that construct would render stable transcripts it would show that also these bases are insignificant, as far as affecting transcript stability goes. Then it could be concluded that only the region between +38 and +47 functions as a cis-acting element.

Another reason for narrowing down the 5`UTR is that a short 5` UTR that includes all elements important to confer stability on transcripts can have many useful traits. For example, the *Chlamydomonas* chloroplast is a popular "machine" for expressing foreign genes since it can be transformed so easily. The short 5` UTR can then be fused onto the 5` end of these foreign genes and stabilize their transcripts. Of course the amount of accumulated transcripts will depend on the genes own promoter's strength.

5.3 Future approaches to identify trans-acting factors relevant for the Chlamydomonas chloroplast gene rbcL.

Once the cis-acting element is identified, the work to isolate and characterize the trans-acting factor(s) begins.
At least eight proteins bind to the entire 5´ UTR of *rbcL (*Hauser *et al.* 1996*)*. Since only the protein(s) affecting transcript stability are of interest, the importance of having narrowed down the 5´ UTR as much as possible without affecting transcript stability comes into play also here. A protein would need its target RNA binding site to consist of at least four to five nucleotides for it to be specific enough. Otherwise the likelihood of the protein binding other places on the RNA would increase. A protein may *bind* to only 4-5 nucleotides, but it will *cover* a larger area than that. Based on this, and the fact that the probable binding area is only fifteen or so nucleotides long, it seems likely that only one protein binds directly to the RNA. If the protein is part of a multi-protein complex, the other proteins will likely bind to the first one.

So the challenge will be to identify the trans-acting factor(s) binding to this area, and to find if this factor(s) is indeed part of a multi-protein complex or not. This can be done in more than one way, both in vitro and in vivo.

For the in vivo assay, one would have to assume that the trans-acting factor is bound to the RNA. The protein(s) can then be permanently fixed to the RNA by cross-linking transformed cells under UV-light. RNA can be isolated from the cells and the proteins can be identified by mass spectrometry by separating the RNA, label it and check the RNA, one fraction at a time, for bound proteins. Many parameters would come in to play, such as light and dark, duration of UV illumination etc.

For in vitro analysis the starting point will be RNA from the construct of interest. The RNA can be synthesized in vitro. Specific vectors designed for in vitro transcription are used. The transcript is then incubated together with an extract of proteins isolated from *Chlamydomonas*. Then, the same procedure as described for the in vivo assay can be followed, isolating the transcript with the protein(s) bound to it for further investigation of the proteins. The in vivo assay is a better method, but more difficult. The in vitro assay is easier but can lead to unspecific binding between proteins, perhaps binding to the protein of interest, thus hindering the binding between that protein and the RNA. There are several possible ways to do this (Information obtained through personal communication with Uwe Klein).
One probable way to identify the trans-acting factor(s) is by chromatography, using an oligo-dT column (Figure 32). The narrowed down version of the 5` UTR would have to be transcribed with an attached poly-A tail. A stretch of for example only Gs should be included between the poly-A tail and the sequence of interest, *e.g.* the proposed cis-acting element, so that the sequence is not too close to the poly-A tail. The reason why only one kind of bases, *e.g.* Gs, should be added is that it decreases the likelihood of proteins binding to such a sequence. The poly-A tail and the stretch of Gs should be fused to the 3´ terminus of the mRNA, thus providing an open, uninterrupted, binding site for trans-acting factors, since such factors will bind closer to the 5` end than the 3´ . The mRNA can then be fixed to the column matrix through interactions between the As and the Ts. An mRNA containing cell extract can then be passed through the column. Since the goal is to obtain highly specific binding between the mRNA and the trans-acting protein, the most important parameter is the salt concentration. A proteins` charge changes with pH, so you want a salt concentration that gives a pH around the protein of interest`s isolelectric point (pI), and thus neutralizes its charge. Since proteins have different pI, various salt concentrations will have

to be tested until only one protein is bound to the mRNA through the cis-acting element, *e.g*. the protein of interest.

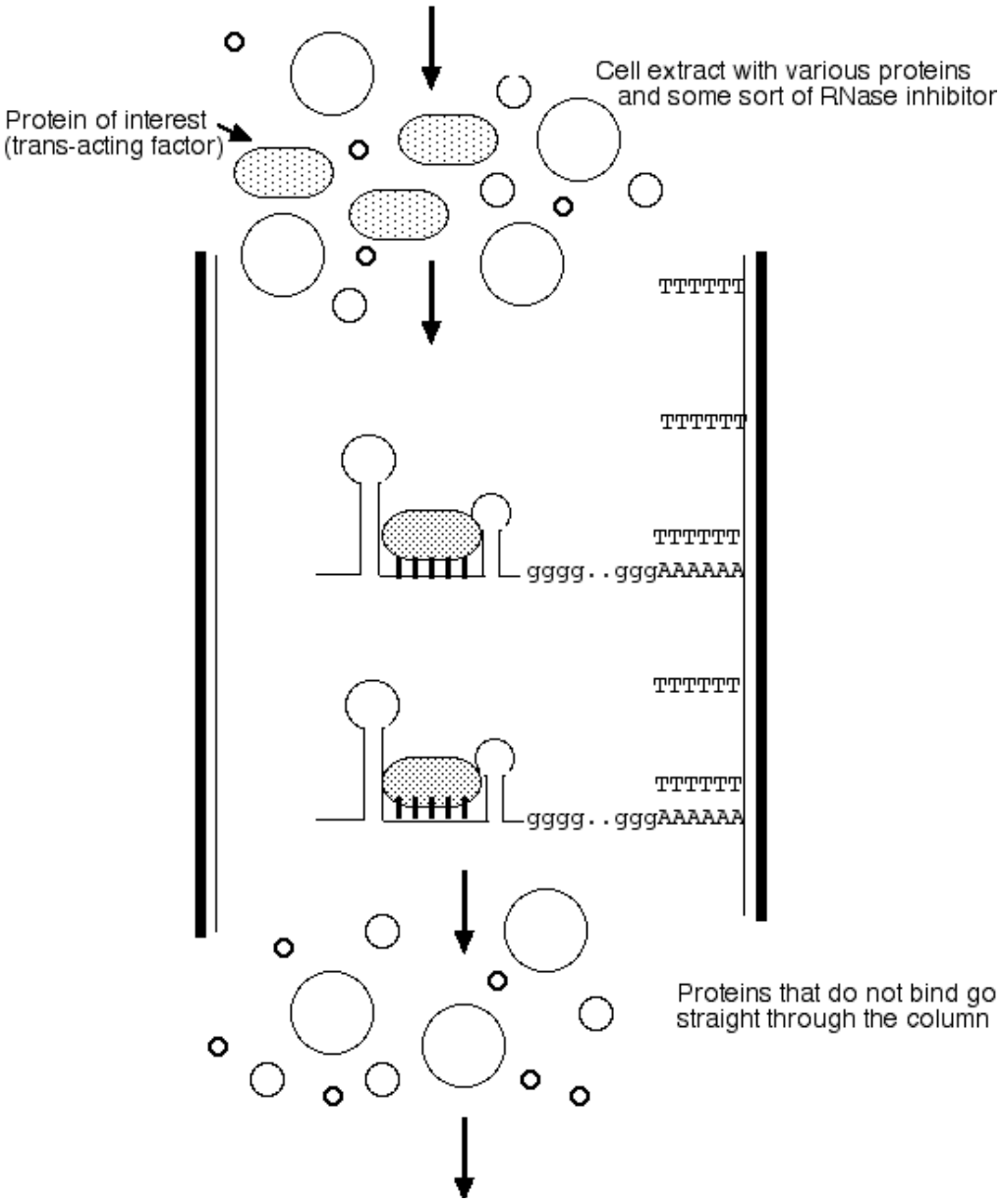

Figure 32: Illustration of interactions between mRNA and a protein of interest. The mRNA is fixed to the column matrix by interactions between the T nucleotides in the column and the poly-A tail attached to the mRNA. When cell extract from Chlamydomonas reinhardtii is passed through the column, the protein of interest, e.g. the trans-acting factor, should interact with the mRNA through specific binding, given that the salt concentration in the column is optimal.

Another parameter that has to be paid attention to is the presence of RNases in a cell extract, since these enzymes break down RNA. Some sort of RNase inhibitor should be added to avoid this.
If a similar mRNA construct, not including the predicted cis-acting element, is mixed with the cell extract previously to passing it through the column, it will function as a control, proving that the binding between the protein and the mRNA is indeed specific, and not due to ionic interactions between any protein and mRNA. Once the cell extract has been passed through the column, and the protein of interest presumably is bound to the mRNA fixed to the matrix, a buffer with a high salt concentration, *e.g*. NaCl, can be passed through the column to elude the protein. The reason why a high salt concentration is used is that the positively charged Na^+ ions and the negatively charged Cl^- ions will compete with the interactions between the mRNA and the protein; Na^+ will bind to the negatively charged mRNA, and the Cl^- will bind to the positively charged protein, thus "dissolving" the original bond between the mRNA and the protein. This way the protein will be eluted as a separate fraction (Information obtained through personal communication with Jon Nissen Meyer).

Trans-acting factors can also be identified by genomic complementation. For this, a true mutant of the gene is required. The mutant would then have to be tested towards each clone in a library until the gene is successfully complemented (Information obtained through personal communication with Uwe Klein).

Once a trans-acting factor(s) is identified, the protein sequence should be obtained and screened for TRP-like domains. If these are present, the trans-acting factor is likely part of a multi-subunit complex (see section 2.1 for details on TRP-like domains) (Boudreau *et al.* 2000).

6 CONCLUSION

Based on the results from this study it can be concluded that deletion of the nucleotides between position +54 and +95, relative to the TSS (+1), in the 5`UTR of the *Chlamydomonas* chloroplast gene *rbcL,* does not affect transcript stability.

It can also be concluded that the nucleotides deleted are not part of a cis-acting element where trans-acting factors will bind, thus the part of the *rbcL* 5' UTR important for transcript stability has been narrowed down further. It now includes, at the most, the nucleotides between positions +38 and +54. Once a deletion of the nucleotides +48 to +54 has been made and its effect on transcript stability has been investigated as in this study, the stability-important part of the 5' UTR can be defined further.

A ribosome binding-site, important for translation, is located in the deleted region. Since this can be deleted without affecting transcript stability, it can also be concluded from this project that there is no interdependent link between transcription and translation for the *rbcL* gene in *Chlamydomonas*.

7 REFERENCES

Adams M (2000). *Chlamydomonas* Teaching Centre http://149.152.32.229/~mikeadams/

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Lewis J, Raff M and Roberts K (2002). Molecular Biology of The Cell - 4th edition (**ISBN** 0815332181). *Garland Science*, New York, USA

Allison L A, Simon L D and Malinga P (1996). Deletion of *rpoB* reveals a second distinct transcription system in plastids of higher plants. *The EMBO Journal*. **15**: 2801-2809

Antonisen I L, Salvador M L and Klein U (2001). Specific sequence elements in the 5′ untranslated regions of *rbcL* and *atpB* gene mRNAs stabilize transcripts in the chloroplast of *Chlamydomonas reinhardtii*. *RNA* **7**: 1024-1033.

Blowers A D, Bogorad L, Shark K B and Sanford J C (1989). Studies on *Chlamydomonas* chloroplast transformation: foreign DNA can be stably maintained in the chromosome. *The Plant Cell*. **1**: 123-132.

Boudreau E N J, Lemaire SD, Ossenbuhl F and Rochaix JD (2000). The *Nac2* gene of *Chlamydomonas* encodes a chloroplast TPR-like protein involved in *psbD* mRNA stability. *The EMBO Journal*. **19**: 3366-3376.

Boynton J E, Gillham N W, Harris E H, Holster J P, Johnson A M, Jones A R, Randolph-Anderson B L, Robertson D, Klein T M, Shark K B and Sanford J C (1988). Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. *Science*. **240**: 1534-1538.

Dellaporta S L, Wood J and Hicks J B (1983). A plant DNA minipreparation: version II. *Plant Molecular Biology*. **1**: 19-21.

Drager R G, Girard-Bascou J, Choquet Y, Kindle K L and Stern D B (1998). *In vivo* evidence for 5′-3′exoribonuclease degradation of an unstable chloroplast mRNA. *The Plant Journal*. **13**: 85-96.

Drapier D, Girard-Bascou J, Stern D B and Wollman F-A (2002). A dominant nuclear mutation in *Chlamydomonas* identifies a factor controlling chloroplast mRNA stability by acting on the coding region of the *atpB* transcript. *The Plant Cell*. **31**: 687-679.

Emory SA, Bouvet P and Belasco JG (1992). A 5'-terminal stem-loop structure can stabilize mRNA in Escherichia coli. *Genes & Development*. **6**: 135-148.

Feinberg A P and Vogelstein B (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry*. **132**: 6-13.

Hajdukiewicz P T J, Allison L A and Maliga P (1997). The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *The EMBO Journal*. **16**: 4041-4048.

Harris E H (1988). The *Chlamydomonas* Sourcebook. A comprehensive Guide to Biology and Laboratory Use (**ISBN:** 978-0123268808). *Academic Press*, San Diego, USA.

Harris E H (2001). *Chlamydomonas* as a Model Organism. *Annual Review of Plant Physiology and Plant Molecular Biology*. **52**: 363-406.

Hartman F C and Harpel M R (1994). Structure, function, regulation, and assembly of D-ribulose-1,5 biphosphate carboxylase/oxygenase. *Annual Review of Biochemistry*. **63**: 197-234.

Haugen P M (2004). Point mutations in the 5' UTR of the *Chlamydomonas reinhardtii* gene rbcL gene destabilize its transcripts. Master thesis from University of Oslo.

Hauser C R, Gillham N W and Boynton J E (1996). Translational regulation of chloroplast genes. Proteins binding to the 5'-untranslated regions of chloroplast mRNAs in *Chlamydomonas reinhardtii*. *Journal of Biological Chemistry*. **271**: 1486–1497.

Hudock G A and Levine R P (1964). Regulation of Photosynthesis in *Chlamydomonas reinhardi*. *Plant Physiology*. **39**: 889–897

Kathir P, LaVoie M, Brazelton W J, Haas N A, Lefebvre P A and Silflow C D (2003). Molecular map of the *Chlamydomonas reinhardtii* Nuclear Genome. *Eukaryotic Cell*. **2**: 362-379.

Klein U, De Camp J D and Bogorad L (1992). Two types of chloroplast gene promoters in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences USA*. **89**: 3453-3457.

Klein U, Salvador M L and Bogorad L (1994). Activity of the *Chlamydomonas* chloroplast *rbcL* gene promoter is enhanced by a remote sequence element. *Proceedings of the National Academy of Sciences USA*. **91**: 10819-10823

Komine Y, Kwong L, Anguera M C, Schuster G and Stern D B (2000). Polyadenylation of three classes of chloroplast RNA in *Chlamydomonas reinhardtii*. *RNA*. **6**: 598-607.

Lister D L, Bateman J M, Purton S and Howe C J (2003). DNA transfer from chloroplast to nucleus is much rarer in *Chlamydomonas* than in tobacco. *Gene*. **316**: 33-38.

Magee A M and Kavanagh T A (2002). Plastid genes transcribed by the nucleus-encoded plastid RNA polymerase show increased transcript accumulation in transgenic plants expressing a chloroplast-localizd phage T7 RNA polymerase. *Journal of Experimental Botany*. **53**: 2341-2349.

Maier R M, Zeltz P, Kossel H, Bonnard G, Gualberto J M and Grienenberger J M (1996). RNA editing in plant mitochondria and chloroplasts. *Plant Molecular Biology*. **32**: 343-365.

Maul J E, Lilly J W, Cui L, dePamphilis C W, Miller W, Harris H E and Stern D B (2002). The *Chlamydomonas reinhardtii* Plastid Chromosome: Islands of Genes in a Sea of Repeats. *The Plant Cell*. **14**: 2659-2679.

Meurer J, Berger A and Westhoff P (1996). A nuclear mutant of *Arabidopsis* with impaired stability on distinct transcripts of the plastid *psbB, psbD/C, ndhH* and *ndhC* operons. *The Plant Cell*. **8**: 1193-1207

Mizohata E, Matsumura H, Okano Y, Kumei M, Takuma H, Onodera J, Shibata N and Kai Y (2002). Crystal structure of activated ribulose-1,5-bisphosphate carboxylase/oxygenase from green alga *Chlamydomonas reinhardtii* complexed with 2-carboxyarabinitol-1,5-bisphosphate. *Journal of Molecular Biology*. **316**: 679-691.

Monde R A, Schuster G and Stern D B (2000). Processing and degradation of chloroplast mRNA. *Biochimie*. **82**: 573-582.

Nickelsen J, Fleischmann M, Boudreau E, Rahire M and Rochaix JD (1999). Identification of cis-acting RNA leader elements required for chloroplast *psbD* gene expression in *Chlamydomonas. The Plant Cell*. **11**: 957-970

Palmer J D (1985). Comparative organization of chloroplast genomes. *Annual Review of Genetics*. **19**: 325-354.

Palmer J D (2003). The symbiotic birth and spread of plastids: how many times and whodunit? *Journal of Phycology*. **39**: 4-11.

Rochaix JD (1996). Post-transcriptional regulation of chloroplast gene expression in *Chlamydomonas reinhardtii. Plant molecular Biology*. **32**: 327-341.

Rochaix JD (2002). The three genomes of *Chlamydomonas*. *Photosynthesis Research*. **73**: 285-93.

Rott R, Drager R G, Stern D B and Schuster G (1996). The 3` untranslated regions of chloroplast genes in *Chlamydomonas reinhardtii* do not serve as efficient transcriptional terminators. *Molecular and General Genetics*. **252**: 676-683.

Rymarquis L A, Higgs D C and Stern D B (2006). Nuclear suppressors define three factors that participate in both 5` and 3` processing of mRNAs in *Chlamydomonas* chloroplasts. *The Plant Journal*. **46**: 448-461.

Sakamoto W, Kindle K L and Stern D B (1992). *In vivo* analysis of *Chlamydomonas* chloroplast *petD* gene expression using stable transformation of β-glucoronidase translation fusions. *Proceedings of the National Academy of Sciences USA*. **90**: 497-501.

Salvador M L, Klein U and Bogorad L (1993). 5` sequences are important positive and negative determinants of the longevity of *Chlamydomonas* chloroplast gene transcripts. *Proceedings of the National Academy of Sciences USA*. **90**: 1556-1560.

Salvador M L and Klein U (1999). "The redox state regulates RNA degradation in the Chloroplast of *Chlamydomonas reinhardtii*." *Plant Physiology*. **121**: 1367-1374.

Salvador M L, Suay L, Anthonisen I L and Klein U (2004). Changes in the 5` untranslated region of the *rbcL* gene accelerate transcript degradation more than 50-fold in the chloroplast of *Chlamydomonas reinhardtii*. *Current Genetics*. **45**: 176-182.

Sambrook J and Russel D W (2001). Molecular Cloning - A Laboratory Manual 1 and 2 (**ISBN**: 978-0879695774). *Cold Spring Harbor Laboratory Press*, New York, USA.

Singh M, Boutanev A, Zucchi P and Bogorad L (2000). Gene elements that affect the longevity of *rbcL* sequence-containing transcripts in *Chlamydomonas reinhardtii* chloroplasts. *Proceedings of the National Academy of Sciences USA*. **98**: 2289-2294.

Spreitzer R J (1993). Genetic dissection of Rubisco - structure and function. *Annual Review of Plant Physiology and Plant Molecular Biology*. **44**: 411-434.

Spreitzer R J (2003). Role of the small subunit in ribulose-1,5-biphosphate carboxylase/oxygenase. *Archives of Biochemistry and Biophysics*. **414** (141-149).

Stampacchia O, Girard-Bascou J, Zanasco J L, Zerges W, Bennoun P and Rochaix JD (1997). A nuclear encoded Function Essential for Translation of the Chloroplast *psaB* mRNA in *Chlamydomonas*. *The Plant Cell*. **9**: 773-782.

Stern D B, Higgs D C and Yang J (1997). Transcription and translation in chloroplasts. *Trends in Plant Science*. **2**: 308-315.

Stern D B and Kindle K L (1993). 3` maturation of the *Chlamydomonas reinhardtii* chloroplast *atpB* mRNA is a two-step process. *Molecular and Cellular Biology*. **13**: 2277-2285.

Suay L, Salvador M L, Abesha E and Klein U (2005). Specific roles of 5` RNA secondary structures in stabilizing transcripts in chloroplasts. *Nucleic Acid Research*. **33**: 4754-4761

Sueoka N (1960). Mitotic replication of deoxyribonucleic acid in *Chlamydomonas* reinhadtii. *Proceedings of the National Academy of Sciences USA*. **46**: 83-91.

Tanaka T and Weisblum B (1975). Construction of a colicin E1-R factor composite plasmid in vitro: means for amplification of deoxyribonucleic acid. *Journal of Bacteriology*. **121**: 354-362.

Vahrenholz C, Riemen G, Pratje E, Dujon B and Michaelis G (1993). Mitochondrial DNA of *Chlamydomonas reinhardtii*: the structure of the ends of the linear 15.8 kb genome suggests mechanisms for DNA replication. *Current Genetics*. **24**: 241-247.

Vajstij F E, Boudreau E, Lemair S D, Goldschmidt-Clermont M and Rochaix JD (2000) Characterization of *mbb1*, a nucleus encoded tetratricopeptide-like repeat protein required for expression of the *psbB/psbT/psbH* gene cluster in *Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences USA*. **97**: 14813-14818

Witsø A (2006). Effect of adding 20 nucleotides to the 5´ terminus of chimeric *rbcL*:GUS transcripts on transcript accumulation in chloroplasts of *Chlamydomonas reinhardtii.* Master thesis from University of Oslo.

Yosef I, Irihimovitch V, Knopf J A, Cohen I, Orr-Dahan I, Nahum E, Keasar C and Shapira M (2004). RNA binding Activity of the ribulose-1,5-biphosphate Carboxylase/Oxygenase large Subunit from *Chlamydomonas reinhardtii*. *The Journal of Biological Chemistry*. **279**: 10148-10156.

Zarins K V (2006). Addition of 10 extra nucleotides to the 5´ UTR of a chimeric *rbcL*: GUS gene destabilizes transcripts in the chloroplast of *Chlamydomonas reinhardtii*. Master thesis from university of Oslo.

8 WEB REFERENCES

Full references to figures taken from websites

Figure 9:
http://images.google.com/imgres?imgurl=http://www.vscht.cz/eds/knihy/uid_es-002/figures/chloroplast.01.jpg&imgrefurl=http://www.vscht.cz/eds/knihy/uid_es-002/motor/index.obrazky.html&h=477&w=600&sz=32&hl=en&start=12&um=1&tbnid=sju90P441HomqM:&tbnh=107&tbnw=135&prev=/images%3Fq%3Dchloroplast%26svnum%3D100%26um%3D1%26hl%3Den%26safe%3Doff%26client%3Dsafari%26rls%3Den%26sa%3DG

Figure 10:
http://images.google.com/imgres?imgurl=http://www.rsbs.anu.edu.au/profiles/Brian_Gunning/Web%2520PCB/Images/Rubisco%2520reactions.jpg&imgrefurl=http://www.rsbs.anu.edu.au/profiles/Brian_Gunning/Web%2520PCB/Images/Rubisco%2520reactions.htm&h=602&w=802&sz=127&hl=en&start=1&um=1&tbnid=Hn3GHM_guJWsFM:&tbnh=107&tbnw=143&prev=/images%3Fq%3Drubisco%26svnum%3D100%26um%3D1%26hl%3Den%26safe%3Doff%26client%3Dsafari%26rls%3Den)

Figure 15:
http://www.fermentas.com/techinfo/nucleicacids/mappbluescriptiiskks.htm

Figure 18:
http://images.google.com/imgres?imgurl=http://fisio.dipbsf.uninsubria.it/scuola/biorad1.jpg&imgrefurl=http://fisio.dipbsf.uninsubria.it/scuola/biolistic.html&h=363&w=363&sz=72&hl=en&start=1&um=1&tbnid=PZXEfDh8l11VyM:&tbnh=121&tbnw=121&prev=/images%3Fq%3DPDS-1000%2BHe%26svnum%3D100%26um%3D1%26hl%3Den%26safe%3Doff%26client%3Dsafari%26rls%3Den